OUTLINE
Current
cover

nature

THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE

ALL TOGETHER NOW

Paris climate talks bring high hopes for a new global agreement PAGE 425

SUSTAINABILITY

SPACESHIP EARTH

Five books that pinpointed our planetary limits

PAGE 443

IMAGING TECHNIQUES

ULTRASOUND IN THE BLOOD

Super-resolution view of circulation in the rat brain

PAGES 451 & 499

METHODOLOGY

ANTIBODY CONFUSION

The search for antibodies you can trust

PAGE 545

NATUREASIA.COM
26 November 2015
Vol. 527, No. 7579

THIS WEEK

EDITORIALS

BRAZIL Legal battle over purported cancer drug tests limits of 'right to try' **p.410**

WORLD VIEW No need for a perfect solution from Paris talks **p.411**

CONSERVATION Mallorcan toad species cured of fungal infection **p.412**

The way forward is through Paris

Leaders must come together on a solid agreement at the United Nations climate conference — and then get to work at home by meeting commitments and finding new ways to reduce emissions.

The world's leaders left a fabulous mess in their wake after making a brief appearance at the last major global climate summit. The 2009 Copenhagen negotiations descended into an angry free-for-all, although one basic idea was agreed: that countries, rich and poor, need to step forward with their own climate solutions. This idea stuck and is now at the heart of the negotiations going into the United Nations Paris Climate Conference, where countries will attempt to forge the first ever fully fledged international climate agreement. *Nature* offers a package of stories and commentaries this week (see nature.com/parisclimate) previewing what many expect to be the biggest step so far towards controlling global greenhouse-gas emissions.

That optimism should not be taken as a sign that all is well. Last year was the warmest on record. This year will be warmer still, with average temperatures expected to reach more than 1 °C above pre-industrial levels. An array of impacts are already being documented around the globe, including melting ice, decreasing crop yields and shifting animal-migration patterns. And yet, despite a quarter of a century of increasingly desperate debate, greenhouse-gas emissions continue to rise.

We know that any deal emerging from the Paris conference will not solve the problem. Even if nations follow through on the climate pledges that have been made so far, global emissions are projected to rise until at least 2030, and temperatures could reach 2 °C above pre-industrial levels as early as 2032. The UN has set the goal of limiting any rise to 2 °C, but even this increase would not protect the world's most vulnerable citizens from rising tides, extreme weather and shifting precipitation patterns.

PLETHORA OF PLEDGES

But there are reasons for optimism. Foremost is the fact that a solid majority of nations, accounting for roughly 91% of global emissions, have submitted climate pledges. Many, including those of all developed countries, feature commitments to curb greenhouse-gas emissions. Others, from a plethora of developing countries, focus on sustainable development and adaptation to the impacts of rising temperatures. Even with financial and technological aid, emissions will continue to rise in these countries as governments seek to lift their people out of poverty.

All told, the world's pledges fall short. But for the first time, governments are moving forward collectively; as David Victor and James Leape point out in their Comment on page 439, that is the first step. Although many countries want to make these commitments binding under international law, they will remain voluntary, at least for now. The US Senate's aversion to international treaties is often blamed, but many countries worry about binding commitments given the difficulty of the economic transition that is required. The 1997 Kyoto Protocol included binding commitments from most developed nations — notably excluding the United States — but many developed countries received a free pass. And there were no real consequences for those that did not live up to their obligations.

The focus now is on building a 'pledge-and-review' system that pushes countries to submit their own national commitments, which are then up for review by other governments and groups. There is some evidence that this 'institutionalized peer pressure' can work: 175 countries have voluntarily submitted pledges so far.

Economic and political momentum is building. Renewable energy is growing faster than anybody projected just a few years ago. The consultancy Bloomberg New Energy Finance has projected that renewables will account for two-thirds of the US$12 trillion that will be invested in electricity generation over the next 25 years. Brazil has made huge progress in reducing deforestation, and the palm-oil industry has committed to reduce deforestation in Indonesia and other countries. The countries of the Organisation for Economic Co-operation and Development agreed on 18 November to restrict financing for coal-fired power plants, and the United Kingdom is weighing up a proposal to shut down all of its coal plants by 2025. In the United States, coal is on the ropes thanks to a combination of regulation and cheap natural gas.

"World leaders must come together and signal the way forward."

In Paris, negotiators must provide a strong framework for reporting and verifying climate pledges. Governments, scientists and environmentalists need solid information about who is doing what. And the agreement should require a five-year review process so that governments can identify ways to go even further at the next major climate summit in 2020. Once everybody is pointed in the right direction, the hope is that human ingenuity will kick in, and the world will discover ways to reduce emissions more quickly.

As reported in our News Feature on page 436, however, limiting the temperature rise to 2 °C will be difficult. Barring premature retirement of much of the existing fossil-fuel infrastructure, the only way to get there will be to overshoot the target and then bring atmospheric carbon dioxide concentrations back down later in the century. Unless engineers figure out a simple way to pull CO_2 out of the atmosphere, this probably means deploying bioenergy at massive scales, capturing the CO_2 that is emitted during energy production and pumping it underground.

One day, governments may decide that measures such as extreme decarbonization are necessary. In the meantime, scientists must investigate the social, political and economic realities ahead and research the consequences of rising emissions, including potentially catastrophic shifts in the climate system.

In Paris next week, world leaders must come together and signal the way forward for their governments, their citizens and for businesses and investors. If humans want to keep living on a planet that looks, feels and functions like the one we live on now, it is time to sign an agreement and get to work. ■

Built on trust

Written agreements between parties in research collaborations are not a sign of a lack of faith.

A scuffle that has riled the Chinese scientific community could have been avoided if the parties involved had hammered out the details of their collaboration beforehand.

On 16 November, researchers at Peking University in Beijing claimed discovery of a biological-compass mechanism that could explain how some animals sense magnetism (S. Qin *et al. Nature Mater.* http://doi.org/89v; 2015). But some of the paper's thunder was stolen by a researcher at Tsinghua University, also in Beijing, who reported in September how the same mechanism could be used to manipulate neurons in worms (X. Long *et al. Sci. Bull.* http://doi.org/883; 2015).

When the September paper was published, the lead Peking University researcher cried foul, claiming that his Tsinghua colleague had agreed not to publish until the *Nature Materials* paper came out (see *Nature* http://doi.org/9gg; 2015). University administrators got involved, the Tsinghua researcher was fired, and his graduate student, whose career has been upended, circulated a plea for support to China's scientific community. The Peking researcher has called for his rival's paper to be retracted. Both parties have mustered e-mails and other correspondence to show that the facts are on their side.

A detailed, formalized agreement could have prevented this. When embarking on a collaboration, it can be hard to ask a scientific peer to sign a contract. Lawyers get involved, making it cumbersome and costly. Fencing off rights to patents, authorship, publication and decision-making authority can be tedious and can cause tension. A simple handshake is much more comfortable.

This is true for researchers around the world. But in China, where people are finely tuned to what might make them lose face, the bar is especially high. Asking someone to sign such an agreement feels equivalent to saying that you don't trust them.

A survey of Chinese researchers undertaken by Nature Publishing Group supports that observation. Scientists who had worked abroad were asked about the differences in the working environment in China compared with that in other countries, including the ease of carrying out collaborations. Some noted that Chinese researchers usually do not ask for formal agreements. The reason might be cultural, but it could also be that most universities and research organizations in China do not have the personnel to support this function.

"The biggest hindrance to harmonious collaboration was tension over authorship."

The survey results appear in a 26 November report, *Turning Point: Chinese Science in Transition* (see go.nature.com/ybsatt and go.nature.com/fdwacj; in Chinese). The biggest hindrance to harmonious collaboration, according to interviewees, was tension over authorship — a factor that plays a substantial part in the dispute over the biological-compass papers. In China, assessors of a researcher's achievements focus on papers in which the individual is first or corresponding author. The report suggests that research assessment should take a more balanced approach, and that policymakers can iron out some of these wrinkles.

It is clear that university administrators can help collaborations by providing personnel to deal with the legal aspects. It might be a burden in the short term, but in the long term it would encourage collaboration. Scientists with valuable knowledge who want to protect their rights to priority in publication, patents and other areas deserve as much. ■

Drugs on demand

Controversy in Brazil over access to a purported cancer cure could set a harmful precedent.

A furious debate that is raging in Brazil pits the nation's largest university against hundreds of cancer patients who want access to a compound that some have branded a miracle cure.

But whether the compound holds any benefits at all remains to be seen: it has never been evaluated in human trials. The conflict is an extreme version of a debate that has gone on in the United States and elsewhere, as terminally ill people whose diseases have withstood modern medicine's proved arsenal have demanded access to untested treatments.

As we report on page 420, courts in Brazil have previously sympathized with those demands, ordering the University of São Paulo to provide a compound called phosphoethanolamine to hundreds of patients. People on both sides of this debate are armed with good intentions. The university argues that the drug is untested, and should not be used to give false hope — and unknown side effects — to vulnerable patients. On the other side, it is understandable that people with little hope may prefer the uncertainty of an untested drug to the certainty of a terminal illness.

But there are also concerning reports that some people with cancer are not taking their prescribed medications, for fear that scientifically proven medicine may interfere with the supposed miracle of phosphoethanolamine. The tenor of the debate has also been harmful at times, with some phosphoethanolamine advocates accusing the government or the pharmaceutical industry of actively suppressing further development of the drug.

The sad truth is that the drug is unlikely to be a miracle. In the United States, for example, only one in ten drugs that make it to phase I clinical trials are destined to gain approval from the US Food and Drug Administration (FDA). And phosphoethanolamine has not even made it that far: its promise is backed up by a few publications based on lab and animal tests.

Even so, terminally ill patients may be willing to try a treatment with only the slimmest odds of success. In the United States, several states have passed laws that, to varying degrees, grant such patients the right to try experimental drugs outside the purview of the FDA. The laws have triggered debates of their own, and have come under fire for offering false hope and for potentially leading patients away from other, more promising avenues.

The situation in Brazil is more extreme. A university laboratory is neither a pharmaceutical plant nor a pharmacy; it is not required to follow good manufacturing protocols. There is no oversight to certify what is going into the blue-and-white phosphoethanolamine capsules produced at the University of São Paulo. Neither the compound's side effects nor its efficacy are systematically monitored. To order a university to supply a drug is to show a disregard for the importance of all these safety measures.

The hope of phosphoethanolamine lies in further research. Federal funders in Brazil have said that they will support further preclinical studies of the drug. Researchers are pursuing options for moving the compound into clinical trials, should those animal studies succeed; patients who are interested in pursuing phosphoethanolamine treatment could enrol in the clinical tests. In the meantime, the courts should liberate patients from the legal tug-of-war and uphold the latest decision to halt distribution of phosphoethanolamine until its potential is better understood. ■

➲ NATURE.COM
To comment online, click on Editorials at:
go.nature.com/xhunqv

M. AXELSSON/AZOTE

A 'perfect' agreement in Paris is not essential

Success at the latest climate talks will be a recognition by the world's nations that incremental change will not do the job, says **Johan Rockström**.

So here we go again. Nations are meeting in Paris for their twenty-first attempt to agree on decisive action to avoid what the United Nations defines as dangerous climate change.

The climate negotiations have set this danger threshold at 1.5–2 °C of global warming above pre-industrial levels. With such a guard rail established, the required components of a 'successful' climate deal more or less fall into place. A reasonable chance of attaining 2 °C translates to a finite global carbon budget of about 900 gigatonnes of carbon dioxide from 2015 onward that must be shared in a fair way between all nations.

Can and should the Paris talks deliver an agreement that gives a binding commitment from all nations to meet this outcome? The last time the world gathered for a decisive global agreement on climate change, in Copenhagen in 2009, the remit was that, yes, world leaders needed to do nothing less than decide on a global, legally binding agreement that met the scientific targets of a safe and just future below 2 °C.

But since Copenhagen, the global discourse has changed. In 2009, it was possible to show convincingly only that we needed to tackle the climate challenge; it was not easy to show that it was possible. Today, the need is more apparent than ever. And, more importantly, there is ample evidence that scaling up economically competitive, clean-energy solutions is possible.

Before Copenhagen, economists generally thought that a high oil price was the best way to enable a transition to a decarbonized future. The surprising reality is that low oil prices seem to be the most effective way of ensuring a transition away from fossil fuels. Renewable energy systems compete even at low oil prices, which in turn closes the door on unconventional, expensive oil, such as offshore oil and exploitation in difficult environments such as the Arctic. It also opens a unique window to introducing a global price on carbon — clearly the most effective policy measure for accelerating the transition to fossil-fuel-free energy.

Experience across industrial sectors shows that new solutions can scale up and become part of the mainstream in markets and societies only once they have penetrated at least 15–20% of the marketplace or society. For renewable energy, this penetration has been achieved in enough countries only in the past three to four years.

In this new situation, is it possible to envisage a transformation to a decarbonized world by around 2050 even if Paris does not deliver the 'perfect' agreement? The answer is yes. To get there, the threshold for success in Paris should not be at the level of 'resolving the climate problem' through incremental change, but rather 'the assurance that the world is serious about a transformation'. We need an agreement that is good enough to tip the world decisively towards rapid decarbonization. A new treaty does not need to force nations into compliance, but rather should create confidence and send the right signal — to investors, businesses and societies at large — that the global political leadership is turning irrevocably towards a new sustainable era.

How ambitious must the Paris agreement be to decisively support such a trajectory? To meet the 2 °C limit, the world must cut carbon emissions at about 6% per year. National pledges on the table at Paris will not get us close. From experience, we know that emissions cuts in the range of 0–2% per year are within the realm of incremental policy measures. A range of 2–3% requires ambitious adaptation. Once levels exceeding 3–4% are reached, experience indicates that radical measures are needed, such as carbon taxes and the phasing out of coal power.

These are the kinds of changes needed to decarbonize the world economy, and above all, to send clear signals of a shift from incremental to transformative change. Success in Paris should thus be viewed as an agreement that corresponds to a pace of emissions cuts of greater than 3–4% per year, starting in the 2015–20 window.

In turn, this would suggest that Paris must accumulate 80% of the national pledges needed to stay within the 2 °C guard rail, with at least 20% of the countries committing to more than 4% average cuts per year, to create a large enough critical mass of nations committed to decarbonization and to influence the global logic (see go.nature.com/1uxlyn). Achieving this goal is ambitious but realistic. And it comes with a decent chance that, once nations realize the benefits of decarbonization, they will increase their pledges. It is crucial, therefore, that the Paris agreement allows for recurrent recalibration of the pledges, at least every third or fifth year.

It would be dangerous to allow 'success' to be reduced to a low level of political achievement so that the world continues along an incremental policy path that stands no chance of supporting a transition to decarbonization. Equally, scientists can no longer dismiss as failure an agreement that is not fully in line with the demands of climate science. For if Paris is widely perceived to have failed, political leadership is likely once again to enter a post-Copenhagen climate trauma and instead focus on other more urgent (and politically rewarding) issues. ■

WE NEED AN **AGREEMENT** THAT IS **GOOD ENOUGH** TO TIP THE WORLD **DECISIVELY** TOWARDS RAPID DECARBONIZATION.

Johan Rockström *is chair of the Earth League and director of the Stockholm Resilience Centre.*
e-mail: johan.rockstrom@su.se

RESEARCH HIGHLIGHTS

Selections from the scientific literature

METABOLISM

Gastric surgery alters sweet tooth

Some weight-loss surgeries can diminish cravings for sweets by altering the brain's response to the neurotransmitter dopamine.

Ivan de Araujo of Yale University in New Haven, Connecticut, and his colleagues studied the effects of a duodenal–jejunal bypass, which reroutes food from the stomach directly into the middle part of the small intestine. They found that well-fed mice that did not have the surgery consumed more sugar after previous repeated exposure to sweets. Mice that had the surgery did not develop the same sweet tooth.

Sugar consumption led to the release of dopamine, which is involved in reward responses, particularly when the sugar was administered to the upper region of the intestines (the area bypassed in the surgery). Activating dopamine-sensing neurons restored the sweet cravings in mice that had undergone the surgery.

Cell Metab. **http://doi.org/9dm (2015)**

ALAN CRESSLER

ZOOLOGY

Mollusc sees with its shell

A marine mollusc has hundreds of eyes in its armour that can see images.

Christine Ortiz at the Massachusetts Institute of Technology in Cambridge and her colleagues studied the structural, optical and mechanical properties of the eyes of *Acanthopleura granulata* (pictured) using various experimental and computational techniques. Unlike in most animals, the microscopic lenses are not organic, but are made of the mineral aragonite. These minimize light scattering because they are made of large and aligned crystals. Projecting images through the lenses showed that they could resolve an image of a potential predator of around 20 centimetres in size from about 2 metres away.

The shells are much weaker at these points than elsewhere, but the organism has evolved ways to compensate for the structural weakness, the team found.

Science 350, **952–956 (2015)**

ECOLOGY

Toads saved from killer fungus

Biologists have rid a wild toad species of a lethal fungal disease that threatens amphibians around the world.

The chytrid fungus *Batrachochytrium dendrobatidis* has wiped out many species of frogs and toads. Jaime Bosch at Spain's National Museum of Natural History in Madrid and his team removed tadpoles of the midwife toad (*Alytes muletensis*; **pictured**) from ponds on the Spanish island of Mallorca and treated them in the lab with a drug that kills the fungus. They also drained the ponds and sprayed them with a disinfectant before returning the tadpoles. The fungus disappeared in four out of five treated ponds for two years.

The method may work only in some habitats, the authors say.

Biol. Lett. 11, **20150874 (2015)**

CHRIS MATTISON/NATUREPL.COM

HYDROLOGY

Snow-fed water supply threatened

The southwestern United States, the Iberian Peninsula and parts of the Middle East and other regions are at risk of seasonal water shortages resulting from decreasing snowfall in a warming climate.

Justin Mankin at Columbia University in New York and his colleagues looked at projections from various climate models to determine how warming might affect snowfall and river run-off in more than 400 large basins in the Northern Hemisphere. The team identified a dozen or so snow-sensitive basins that, across all climate models, face an 80–100% risk of declining water supply in the coming decades. Each of the sensitive basins has a current population of more than 1 million people — including the Rio Grande basin spanning Texas and Mexico, the Ebro–Duero

basin in Spain and the Asi basin in Lebanon and Syria.
Environ. Res. Lett. 10, **114016 (2015)**

NUTRITION

Personalized diets for health

People who eat identical meals display different blood glucose levels afterwards, thanks in part to differences in their gut microbes.

Large spikes in blood glucose after eating increase the risk of type 2 diabetes, so dietary guidelines rank foods based on their glycaemic index — an indicator of their effects on blood glucose. Eran Elinav and Eran Segal of the Weizmann Institute of Science in Rehovot, Israel, and their colleagues continuously monitored the diets and lifestyles of 800 people over a week, and found that meals with the same glycaemic index caused widely different glucose levels in participants. By analysing data on the participants' gut microbiomes, physical activity and other clinical factors, the team created personalized diets for 26 people and found that these resulted in lower glucose levels after meals than did non-personalized diets.

The study could partly explain the limited efficacy of universal dietary guidelines, the authors say.
Cell 163, **1079–1094 (2015)**

QUANTUM METROLOGY

Lasers reveal quantum jitters

Ultrafast laser pulses can be used to detect the motion of a single atom, from energetic wiggles to quantum jitters.

Kale Johnson at the University of Maryland in College Park and his colleagues trapped ions of ytterbium and zapped them with laser pulses just 10 picoseconds long. The pulses gave the atom small kicks in momentum of different magnitudes, depending on its internal state. This resulted in a new state that encoded the atom's original motion. After another sequence of pulses, the researchers observed fluorescent light from the atom that allowed them to measure its quantum motion.

The technique could be useful for future quantum computers built from trapped ions, the team says.
Phys. Rev. Lett. 115, **213001 (2015)**

BIOELECTRONICS

Flower given digital power

Researchers have incorporated electronic circuitry into the tissues of a rose.

Magnus Berggren at Linköping University in Norrköping, Sweden, and his colleagues submerged the cut end of a rose stem into a water-based solution of PEDOT, a conducting polymer that is used in printable electronics. Capillary action pulled the polymer up into the rose's vascular tissue, where it came out of solution and self-assembled into wires, some as long as 10 centimetres. By attaching gold probes coated with PEDOT to the wires, the researchers made individual transistors and demonstrated a simple digital circuit. The transistors' electrical performance was on a par with that of conventional printed PEDOT circuits.

The technology could eventually be used to record or regulate plant physiology, the authors say.
Sci. Adv. 1, **e1501136 (2015)**

AGRICULTURAL ECOLOGY

Complex effects of pesticides on bees

Honeybee colonies could be compensating for the harmful effects of certain pesticides by producing more workers, at least in the short term.

Some European countries banned neonicotinoid pesticides in 2013, but this remains controversial because field studies have failed to confirm the adverse effects reported for bees in the lab. Mickaël Henry at the French National Institute of Agricultural Research in Avignon and his colleagues positioned honeybee colonies in farmers' fields so that they were exposed to varying levels of the pesticide thiamethoxam. The team radio-tagged and monitored nearly 7,000 bees, and found that pesticide exposure caused an acceleration in death rate over time.

The colonies, however, compensated for dead foragers by producing more workers and fewer drones. This maintains honey production but could decrease bee reproduction in the long term. The risks of pesticides in the field may be best understood by studying entire colony cycles, the authors say.
Proc. R. Soc. B 282, **20152110 (2015)**

SOCIAL SELECTION

Popular topics on social media

Text-mining block prompts response

A scientist who mines the text of research publications was blocked by the scientific publisher Elsevier from downloading large numbers of its papers — a move that he described in a blog post that was shared by many on social media. Chris Hartgerink, a statistician at Tilburg University in the Netherlands, says that the publisher is hindering his research. Elsevier allows text-mining through the use of a specific application programming interface, and says that this prevents its website from being slowed down by researchers who download large amounts of data. Frank Huysmans, a library scientist at the University of Amsterdam, linked to the blog post on Twitter: "How signing away copyright to academic publishers obstructs content mining research … Strong case for #openaccess #tdm."

NATURE.COM
For more on popular papers:
go.nature.com/pkmi9d

NASA

PLANETARY SCIENCE

Martian moon will break apart

Phobos, one of Mars's two moons, will disintegrate some 20 million to 40 million years from now, and its particles will form the only planetary ring in the inner Solar System.

Benjamin Black and Tushar Mittal of the University of California, Berkeley, made these predictions by analysing tidal and other forces that are currently pulling Phobos (**pictured**) towards Mars. Using a geological model of how rock holds together, they calculated that the moon would rip apart before it smashed into the planet. The resulting ring would be stable for 1 million to 100 million years, they say.
Nature Geosci. **http://dx.doi.org/10.1038/ngeo2583 (2015)**

NATURE.COM
For the latest research published by *Nature* visit:
www.nature.com/latestresearch

SEVEN DAYS *The news in brief*

EVENTS

Whalers fined

An Australian court fined a Japanese company Aus$1 million (US$724,000) on 18 November and found the firm to be in contempt of court for killing minke whales in an area declared a sanctuary by Australia. According to the animal-protection organization Humane Society International (HSI), which, along with the Environmental Defender's Office, brought the case against the firm Kyodo Senpaku Kaisha, this is one of the largest fines issued under Australian conservation law. The company caught whales in four different years, despite a 2008 injunction against the practice, says the HSI.

Climate repeals

The US Senate voted on 17 November to repeal a pair of regulations by the Environmental Protection Agency that would limit carbon emissions from new and existing power plants. Votes on both rules were led by Republicans and passed by a margin of 52–46; the House of Representatives is considering similar resolutions. Coming just two weeks before the United Nations climate summit in Paris, the resolutions are largely symbolic. US President Barack Obama promised to veto both repeals, and supporters do not have the two-thirds majority needed to override a veto.

Statoil Arctic exit

Norwegian energy company Statoil announced on 17 November that it would cease exploration for gas and oil in Alaska's Chukchi Sea. The decision comes just over a month after Royal Dutch Shell suspended its own exploration off the Alaskan coast, citing regulatory uncertainty and a disappointing survey of the area's fossil-fuel prospects. The Statoil decision sees the company exit early from 16 leases that were set to expire in 2020.

JASON REED/REUTERS

Tasmanian devils returned to the wild

Tasmania has 39 more wild devils, after the latest batch of healthy individuals was released from the Devils Ark Santuary (pictured is manager Dean Reid) onto the Forestier Peninsula on 18 November. The area was cleared of Tasmanian devils (*Sarcophilus harrisii*) after an infectious cancer that is devastating populations of the endangered animals was detected there in 2004. A 'devil-proof fence' has now been installed to prevent the new, healthy population from mixing with animals afflicted with the deadly and infectious devil facial tumour disease.

L'Aquila verdict

Italy's highest court of appeal on 20 November upheld a decision to acquit 6 seismologists accused of manslaughter in regard to the 2009 L'Aquila earthquake, which killed more than 300 people. Prosecutors claimed that the scientists misled townspeople about the risk, leading them to stay in their homes instead of seeking safety. The scientists were originally given six-year prison sentences, but an appeals court in L'Aquila acquitted them last November, and reduced to two years the sentence of Bernardo De Bernardinis, former deputy director of the Italian Civil Protection Department, who was also convicted. De Bernardinis's reduced sentence was upheld; he still faces a separate charge of manslaughter.

Ebola setback

In a setback to efforts to end the Ebola epidemic in West Africa, the World Health Organization announced three new cases of the disease in Liberia on 20 November. One of those individuals, a 15-year-old boy, died on 23 November. The country had been declared Ebola-free on 3 September. Sierra Leone was declared Ebola-free on 7 November, and the last case in Guinea was reported on 29 October, leading to hopes that the epidemic, which began in December 2013, might finally be nearing an end.

Pandemic report

A panel of physicians, scientists and policy experts has called for major reforms to the World Health Organization and other international health-response systems following the Ebola epidemic that has killed more than 11,000 people. The panel, convened by Harvard University in Cambridge, Massachusetts, and the London School of Hygiene and Tropical Medicine, released its report on 22 November (S. Moon *et. al. Lancet* http://doi.org/9gf; 2015). It also recommends measures to improve prevention, detection

and response to outbreaks, and to speed research on diseases that cause them. See go.nature.com/jxxvs6 for more.

Rare rhino dies

Northern white rhinoceroses (*Ceratotherium simum cottoni*) are one step closer to extinction, after a 41-year-old female named Nola had to be put down after surgery at the San Diego Zoo Safari Park in California on 22 November. The last three remaining individuals — two females that cannot reproduce naturally and a male with a low sperm count — live at Ol Pejeta Conservancy in Kenya. Conservationists hope that the species can be saved through assisted reproduction techniques, using southern white rhinos (*Ceratotherium simum simum*) as surrogates.

KRISTOPHER GRUNERT/CORBIS

POLICY

CRISPR cress

The Swedish Board of Agriculture on 17 November told two Swedish universities that they do not need special approval for field trials of some cress (*Arabidopsis*, **pictured**) varieties mutated by the CRISPR–Cas9 gene-editing technique. In June, the European Commission had asked European Union member states to hold back on such rulings until it makes its own proposals on how to regulate organisms modified by new genetic techniques. But the Swedish authority said decisions needed to be made now, so that trials can be prepared for the next growing season.

Chimps retired

The US National Institutes of Health (NIH) is ceasing its chimpanzee-research programme altogether, two years after retiring most of its chimps. In a 16 November e-mail to the agency's administrators, NIH director Francis Collins announced that the 50 NIH-owned animals that remain available for research will be sent to sanctuaries. The agency will also develop a plan to phase out NIH support for the remaining chimps that are supported, but not owned, by the NIH. See page 422 for more.

Coal curbs

The Organisation for Economic Co-operation and Development agreed on 18 November to restrict public financing for coal-fired power plants. Two years in the making, the agreement removes support for large, low-efficiency coal-fired plants while maintaining support for medium-sized, high-efficiency plants in countries facing energy shortages, and for small, less-efficient plants in the poorest countries. The restrictions will not apply to any coal-fired plants that are equipped to capture and store carbon emissions.

FUNDING

UK research review

A tensely awaited report into the future of the major UK research-funding agencies, released on 19 November, suggests the creation of a powerful umbrella organization called Research UK to manage the agencies. The review was led by Nobel-prizewinning geneticist Paul Nurse. Many scientists feared that it would recommend a total merger of the research councils, which collectively distribute some £3 billion (US$4.6 billion) of government research funding each year. Nurse recommends that Research UK be led by an experienced researcher, who would in effect be boss of the heads of the seven discipline-based councils. See go.nature.com/2rwzeu for more.

COMING UP

30 NOVEMBER
The leaders of the world's nations gather to broker a climate deal at the United Nations' Paris Climate Change Conference.
nature.com/parisclimate

1–3 DECEMBER
Washington DC hosts the International Summit on Human Gene Editing.
go.nature.com/huzip3

2 DECEMBER
The European Space Agency's LISA Pathfinder satellite, which will hunt for gravitational waves, launches from Kourou, French Guiana.
go.nature.com/rxrzuc

BUSINESS

Mega-merger

Two major pharmaceutical firms are to merge in a US$160-billion deal, they announced on 23 November. Pfizer of New York will combine with Allergan, based in Dublin, in a merger that is expected to be completed by the end of 2016. The resulting firm will be named Pfizer but will be headquartered in Dublin — providing a significant tax break for the US firm — and will have more than 100 medicines in mid-to-late-stage development.

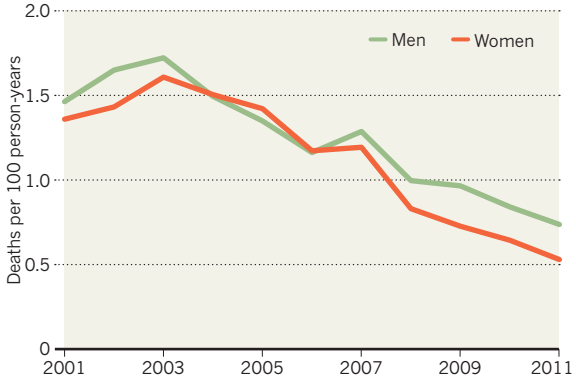
NATURE.COM
For daily news updates see:
www.nature.com/news

TREND WATCH

The availability of antiretroviral drugs for HIV has increased women's lifespans more than men's in KwaZulu Natal in South Africa, concludes a study of more than 98,000 people (J. Bor *et al. PLoS Med.* **12,** e1001905; 2015). Since free antiretroviral treatment became available in South Africa in 2004, declines in life expectancy have reversed for both genders. But progress is uneven, with women gaining more years of life than men. The authors recommend that HIV outreach activities be targeted to men.

HIV LIFE EXPECTANCIES DIFFER BY GENDER

Freely available antiretroviral therapy has decreased the mortality rate of HIV-positive women more than are HIV-positive men in one region of rural South Africa.



SOURCE: J. BOR *ET AL. PLOS MED.* **12,** E1001905 (2015)

NEWS IN FOCUS

PAUL DARROW/NYT/REDUX/EYEVINE

AquAdvantage Atlantic salmon (at back) grow to twice the size of an normal Atlantic salmon (*Salmo salar*) over the same time.

BIOTECHNOLOGY

Transgenic salmon leaps to the dinner table

Long-awaited decision by US government authorizes the first genetically engineered animal to be sold as food.

BY HEIDI LEDFORD

A breed of fast-growing Atlantic salmon rocketed to celebrity status on 19 November when it became the first genetically engineered animal to be approved for human consumption in the United States.

The landmark decision by the US Food and Drug Administration (FDA) releases the 'AquAdvantage' salmon from two decades of regulatory limbo — but it could also revitalize an industry that has waited a long time for any sign that its products might make it to market.

"It opens up the possibility of harnessing this technology," says Alison Van Eenennaam, an animal geneticist at the University of California, Davis. "The regulatory roadblock had really been disincentivizing the world from using it."

The FDA decision comes at a time when the US government is re-evaluating how it regulates genetically engineered crops and animals. On 2 July, the White House Office of Science and Technology Policy said that it will update those regulations — for the first time since 1992 — over the next year. And at a meeting on 18 November, the US Department of Agriculture (USDA) discussed preliminary plans to revise its guidelines for genetically engineered crops.

A key driving force for these discussions is the recognition that current regulations may not cover crops and animals engineered using cutting-edge techniques, such as the ▶

CRISPR–Cas9 system, that allow researchers to make targeted changes to the genome. The USDA has already determined that its regulations do not apply to several genome-edited crops. Van Eenennaam says that it is still unclear how the FDA will regulate animals that have been engineered using that technology.

"There is a lot going on these days," says Greg Jaffe, director of biotechnology at the Center for Science in the Public Interest in Washington DC. "But obviously, up until the decision about the salmon, people were mostly focusing on the crop side."

AquaBounty Technologies, based in Maynard, Massachusetts, filed its first application to the FDA for approval of the salmon in 1995. The agency completed its food-safety assessment in 2010, and released its environmental-impact statement at the end of 2012. The long delay between the completion of those steps and a final decision led to rumours of political interference.

But Laura Epstein, a senior policy analyst for the FDA's Center for Veterinary Medicine, says that the approval took so long because it was the first of its kind. "With most products that are the first of its kind, we are very careful," she says. The agency also had to wade through many public comments before it could issue a decision, she adds.

It is unclear how the salmon will fare on the market. AquAdvantage fish produce extra growth hormone, allowing them to grow to market size in 18 months, rather than the usual 3 years. In the time since AquaBounty first filed for approval, fisheries have bred conventional salmon that grow just as fast, says Scott Fahrenkrug, chief executive of Recombinetics, an animal-biotechnology firm in St Paul, Minnesota.

"It opens up the possibility of harnessing this technology."

Then there is the matter of consumer acceptance: several grocery chains have said that they will not carry the salmon, which, even at full production, would amount to only a tiny fraction of total US salmon imports. "It's a drop in the bucket," says Jaffe. "Consumers would have to hunt to find salmon that are genetically engineered, as opposed to avoiding them."

Still, the FDA's approval met with swift opposition from some environmental and food-safety groups. Although AquaBounty uses physical and biological safeguards to reduce the chance that its salmon will escape into the wild, opponents fear that an accidental release could alter natural ecosystems. They are also unhappy that the FDA will allow the fish to be sold without any label to indicate that it is genetically engineered.

"Huge numbers of people have said, 'Yes, we want it labelled'," says Jaydee Hanson, a senior policy analyst at the Center for Food Safety, an environmental-advocacy group in Washington DC. "If this is such a good product, the company itself should be saying it will label it."

The FDA declined to comment on whether other applications for genetically engineered animals are in the regulatory pipeline. But Fahrenkrug says that his company is developing several such animals, including cattle that do not have to be dehorned and pigs that do not need to be castrated.

Recombinetics' animals are engineered using genome-editing techniques that Fahrenkrug argues do not require FDA approval. The agency regulates animals that are engineered using a "recombinant DNA construct", but his animals are modified by injecting protein and RNA into embryos. "It's a treatment, not a transgene," he says.

The FDA has yet to announce how it will view such animals, but Fahrenkrug takes approval of the salmon as a sign that the agency is willing to allow them onto the market. "I'm feeling optimistic now," he says. ■

PARIS CLIMATE TALKS

Pledges raise hopes ahead of climate talks

Momentum builds for a new treaty as world leaders prepare to descend on Paris.

BY JEFF TOLLEFSON

The road to a new global climate treaty has been slow and plodding. But years of delicate negotiations have given way to cautious optimism as more than 190 nations prepare for the marathon climate talks that begin in Paris on 30 November.

Some long-running disputes remain, such as the debate about what cuts in greenhouse-gas emissions can be expected of developing nations compared with their developed counterparts. But there are many signs that the summit, convened by the United Nations, will succeed in crafting a global climate agreement. These include significant commitments by several major players, including the United States and China, to reduce emissions of greenhouse gases.

"We are in for some tense negotiations, but I think we'll come out of the other end with an agreement," says Saleemul Huq, director of the International Centre for Climate Change and Development in Dhaka, Bangladesh, and adviser to a negotiating bloc of the least-developed countries.

And although Paris is still reeling from the deadly terror attacks of 13 November, which led the authorities to increase security for the meeting and cancel a big climate march, more than 130 heads of government and state are still expected to attend the two-week summit.

The last major push for a climate treaty faltered in Copenhagen six years ago over whether developing countries should be asked to match developed countries and make voluntary commitments to reduce emissions. The political situation has evolved since then and more than 165 countries have submitted pledges to combat climate change. Although these pledges would not cut greenhouse-gas emissions enough to meet the UN goal of limiting global warming to 2 °C above pre-industrial levels, they show a level of commitment that was missing in Copenhagen.

"Countries are bringing more political will than ever before, and so we'll see if the process can deliver," says Elliot Diringer, executive vice-president of the Center for Climate and Energy Solutions, an environmental think tank in Arlington, Virginia. "This agreement has the potential to be a significant turning point."

Despite a lingering — and potentially volatile — debate about whether those commitments will be legally binding under international law, they are expected to remain voluntary. One of the biggest obstacles to a binding agreement is the US Senate. On 17 November, Republican senators pushed through legislation seeking to block regulations to limit greenhouse-gas emissions from power plants. US President Barack Obama can veto these bills, but he cannot force the Senate, which has the power to reject or approve treaties, to endorse a climate agreement that includes binding limits on greenhouse-gas emissions.

As a result, much of the debate will centre on creating mechanisms that allow governments — and civil society — to monitor progress, build trust and ensure accountability.

Environmentalists and many governments are pushing for a five-year review period that would begin immediately after the Paris talks end; governments would need to return to the table with new commitments in 2020.

Huq says that this exercise is particularly important for poor and vulnerable countries, which are pushing for a long-term goal of limiting warming to 1.5 °C. The world is likely to cross a landmark threshold, the 1 °C mark, for the first time in 2015, and Huq admits that stabilizing at 1.5 °C would require emissions reductions so drastic as to be politically impossible at this point. But world leaders should acknowledge that even 2 °C of warming comes with significant impacts on the world's poorest citizens, he says. "We know we are not going to get everything we want in Paris, but it's symbolic."

Samantha Smith, leader of environmental group the WWF's Global Climate and Energy Initiative in Oslo, says that the biggest debate in Paris will be over financial aid to help poor countries to reduce their emissions and cope with the impacts of climate change. In 2010, wealthy nations established a Green Climate Fund and committed to increase climate aid to US$100 billion annually by 2020. Developing countries will be looking for details about that commitment and what comes next.

The good news, Smith says, is that the conversation about climate action has changed, not just within the negotiations but among faith groups, the general public and businesses, many of which will make their own voluntary emissions commitments in Paris. But she cautions that a new global treaty is just a first step. "When we walk out of there, we are still going to have a lot of work to do." ■

ENVIRONMENT

Green Climate Fund faces slew of criticism

First tranche of aid projects prompts concern over operations of fund for developing nations.

BY SANJAY KUMAR

Major questions are swirling around the operations of a United Nations fund that is supposed to channel billions of dollars to help developing nations adapt to climate change and slow its pace.

The Green Climate Fund (GCF) was established at UN talks in Cancún, Mexico, five years ago, and developing nations see it as one of their prime hopes for financial assistance in tackling a warming world.

Yet the fund, which is administered by a small team in Incheon, South Korea, is struggling to raise cash from rich nations. And although it approved its first aid commitments on 6 November at a meeting in Livingstone, Zambia, observers say they are concerned that the GCF has cut corners so as to announce handouts before international climate talks in Paris in December.

"We are worried about the fund's social and environmental safeguards, consultation processes, accountability mechanisms and transparency," says Brandon Wu, a policy analyst who focuses on climate finance at the non-governmental organization (NGO) ActionAid in Washington DC and who attended the Zambia meeting.

The Cancún agreement recommended that climate aid total US$100 billion a year by 2020, but the balance between private and public money, and how much of it would flow through the GCF, has not been made clear.

In the world of climate finance, the GCF is a tiny player. If funding for renewable energy and energy-efficiency programmes is included, hundreds of billions of dollars already flow round the globe each year, says the Climate Policy Initiative (CPI), an international think tank. Still, the GCF is the largest international public climate fund.

ESPEN RASMUSSEN/PANOS

Flood barriers in Bangladesh could find support from a United Nations climate fund.

The fund's initial target was to collect $10 billion before it started handing out cash, which it intends to divide equally between mitigation and adaptation projects. By October, it had received pledges of $10.2 billion — which foreign-exchange rate variations have reduced to $9.1 billion. But only $5.83 billion had been formally agreed, and just $852 million had reached the fund's pocket. The United States is the most significant missing name from the list of donor countries: last year it promised $3 billion, but it has yet to sign an agreement to contribute money.

"At this pace we will not be able to do anything much," says Dipak Dasgupta, an economist and India's representative on the 24-person GCF board. The proposals ▶

approved in Zambia — $168 million for eight climate projects — are "small change", he says. The approvals include a wetlands resilience programme in Peru, climate-resilient infrastructure in Bangladesh and a scheme of 'green bonds' to finance sustainable energy ventures in Latin America and the Caribbean, but seven of the schemes will not receive money until they meet further project-specific conditions.

Developed nations may be reluctant to transfer their money to the fund, says Timmons Roberts, who studies climate change and economic development at Brown University in Providence, Rhode Island. "Many developing countries and NGOs believe that the funding should all flow through the GCF," he says. "However, contributor countries have always defended their ability to funnel their funds through channels they control, whether through their own bilateral agencies (like USAID) or through dedicated World Bank funds."

LACK OF TRANSPARENCY

There are also concerns about how the GCF is run, says Wu, who attended the Zambia meeting as a permitted 'civil society observer'. Wu is worried that indigenous communities were not adequately consulted before the approval of $6.2 million for the Peruvian wetlands programme, for example. GCF documents say that a consultation was carried out, but for this and for other projects, the fund has no independent verification of its claims, says Andrea Rodríguez Osuna, who works in Mexico City for the non-profit environmental law organization AIDA and was also present in Zambia.

PARIS CLIMATE TALKS
A *Nature* special issue
nature.com/parisclimate

Nor is the GCF transparent about its processes, Rodríguez Osuna adds. "The fund has no information disclosure policy and no accountability mechanism, yet the board is approving project proposals," she says.

For the eight projects approved at the board meeting, for example, only proposal documents were publicly available (and in the case of two private-sector projects, only a summary). "These are hardly the unbiased sources of information needed to evaluate a project's merits or any potential negative impacts," Wu says. Project reviews made by the fund's board and by an independent technical advisory panel are not publicly released, and GCF officials repeatedly failed to answer questions asked by *Nature* for this article.

For some, another contentious issue is that the GCF is flowing its money mainly through international organizations, such as multilateral or private banks such as the World Bank and Deutsche Bank — rather than sending it directly to institutions in developing countries where the projects are taking place.

The GCF is still new and is seriously understaffed, Rodríguez Osuna adds; and observers hope that their worries are teething problems. Its executive director, Héla Cheikhrouhou, has promised "many more projects under development".

Claims have already been made that rich nations are upscaling public climate funding. But experts say that there is little clarity on whether the cash is new money, or being re-routed from elsewhere, such as from overseas development assistance funds. "Definitions of what constitutes new money haven't been agreed on," says Barbara Buchner, who leads CPI's global finance programme in Venice, Italy.

There is one thing is for certain, Buchner says — total finance for low-carbon energy projects and for adapting to and mitigating climate change is far short of estimates of the need. "We need trillions, not billions," she says. ■

POLICY

Brazilian courts tussle over unproven cancer treatment

Patients demand access to compound despite lack of clinical testing.

BY HEIDI LEDFORD

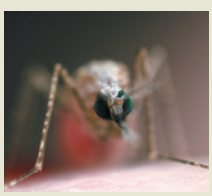
A court in the Brazilian state of São Paulo has cut off distribution of a compound that is hailed by some as a miracle cancer cure — even though it has never been formally tested in humans.

On 11 November, to the relief of many cancer researchers, a state court overturned earlier court orders that had obliged the nation's largest university to provide the compound to hundreds of people with terminal cancer. Although the reversal applies only to requests for the drug by residents of São Paulo state, administrators at the university estimate that it covers about 80% of the orders they have received for the compound.

The compound, phosphoethanolamine, has been shown to kill tumour cells only in lab dishes and in mice (A. K. Ferreira *et al. Anticancer Res.* **32,** 95–104; 2012). Drugs that seem promising in lab and animal studies have a notoriously high failure rate in human trials. Despite this, some chemists at the University of São Paulo's campus in São Carlos have manufactured the compound for years and distributed it to people with cancer. A few of those patients

MORE ONLINE

TOP NEWS



'Gene drive' mosquitoes engineered to fight malaria go.nature.com/r53npf

MORE NEWS

- Graph-theory breakthrough tantalizes mathematicians go.nature.com/8mgjdx
- The world's biggest volcano is a magnetic mix-up go.nature.com/2mqglb
- Ebola experience leaves world no less vulnerable go.nature.com/jxxvs6

PARIS CLIMATE TALKS

Daily news, commentary and video updates from the Paris climate talks parisclimatetalks2015.tumblr.com

SINCLAIR STAMMERS/SPL

NATURE VIDEO

CECILIA BASTOS/USP

Phosphoethanolamine capsules were manufactured at the University of São Paulo.

have claimed remarkable recoveries, perpetuating the compound's reputation as a miracle cure.

Dismayed by this unofficial distribution of phosphoethanolamine, the university's administration moved in September 2015 to shut it down. Patients took the university to court, and in October 2015, Brazil's Supreme Federal Court ruled in favour of one plaintiff who wanted the right to try the compound. A lower court then began granting orders for the university to provide it to others. University officials say that they were soon overwhelmed by more than 800 requests.

"The decision not only ignored the opinion of medical specialists, but also overlooked the fact that the drug has only been tested on animals," says bioethicist Volnei Garrafa at the University of Brasilia. "Such court decisions bring false expectations for patients and their families, creating turmoil in society and confusion between what is safe and what is not."

The Brazilian constitution guarantees universal access to health care, and it is common in Brazil for patients to turn to the courts to access drugs that the state health-care system does not dispense because of their cost, says Garrafa. But phosphoethanolamine presents a different situation, he adds, because it is not really a 'drug' at all. It is not approved by Brazil's National Health Surveillance Agency.

Those who argue that people who are terminally ill have a right to try experimental medicines saw the decision earlier this year as a significant victory. But to the university administration, drug regulators and cancer researchers, it showed blatant disregard for the basic scientific principle that a drug should be demonstrated to be safe and effective before being given to patients outside of a clinical trial.

"It's a violation of the autonomy of the university," says Marco Antonio Zago, a physician and president of the University of São Paulo. "We are seen as a factory to produce something that we do not believe should be done."

Phosphoethanolamine is an important building block of the lipids that make up cell membranes. The compound can also act as a molecular signal that activates certain cellular processes. Although some studies do suggest that the compound may kill cancer cells in isolated cells and mice, it is not entirely clear how the compound brings about this response. Biochemist Durvanei Augusto Maria at the Butantan Institute in São Paulo believes that the compound may be imported into tumour cells and, once inside, trigger processes that cause the cell to self-destruct. Immunologist James Venturini at São Paulo State University and his colleagues have found that phosphoethanolamine may modulate the immune system's response to cancer or affect cell division (M. S. P. de Arruda *et al. Braz. Arch. Biol. Technol.* **54,** 1203–1210; 2011).

But to justify using phosphoethanolamine in people, Venturini says, one would have to rigorously test it in a series of clinical trials using human volunteers. "I strongly believe that double-blind, randomized clinical studies are necessary," he says.

And even before such trials, further preclinical studies would have to be done, says Jailson Bittencourt de Andrade, secretary for research-and-development policy at Brazil's science and technology ministry. The ministry plans to fund those studies, he says, and has already asked several research laboratories in the country to do the work. If those tests and subsequent clinical trials are successful, he says, the ministry will also fund the research needed to scale up phosphoethanolamine production to the quantities and quality needed for an approved drug.

That process will take years. In the meantime, lawyers representing people with cancer have vowed to appeal against the latest ruling. If those appeals succeed, de Andrade worries that people will not wait until all the tests are completed, and may even abandon conventional treatment in favour of phosphoethanolamine. "Many patients have come forward and said they have tried the drug and it has worked for them," he says. "So the other patients and their families — they want phosphoethanolamine now." ■ SEE EDITORIAL P.410

TIMEKEEPING

Leap-second decision delayed

Nations fail to agree on whether to scrap an adjustment that keeps official time in sync with Earth's rotation.

BY ELIZABETH GIBNEY

A leap second is gone in the blink of an eye. But a decision on whether to ditch these occasional time insertions — which keep official time synced with Earth's rotation — has been delayed for at least eight years.

This month, the International Telecommunication Union (ITU), which bears responsibility for defining official Coordinated Universal Time (UTC), was expected to reach a consensus. But representatives who discussed the issue at the World Radiocommunication Conference in Geneva, Switzerland, failed to agree on whether the leap second's costs outweigh its benefits.

Leap seconds, which occur once every few years, are necessary because Earth's rotation is slowing in an unpredictable way. Without them, the time of day when the Sun is at the highest point in the sky would drift by about one minute over about 100 years. However, these extra seconds have to be programmed into electronic systems manually and can upset systems that ▶

depend on accurate timings.

Most countries, including China, the United States and large parts of Europe, favour scrapping the leap second and basing UTC on the continuous tick of atomic clocks.

Official time would slowly move out of sync with Earth's rotation, but — given that it would take thousands of years to accumulate a difference that is greater than the shifts already caused by daylight savings time — many argue that this would cause few problems. "We are already shifted by one hour in summer compared to winter time," says Elisa Felicitas Arias, director of the Time Department at the International Bureau of Weights and Measures (BIPM) in Sèvres, France, who wants to scrap the leap second. "Are we affected because of that?" A correction — perhaps a leap minute or hour — could be added once the drift is appreciable.

A small number of countries however, including Russia and the United Kingdom, want to keep the leap second. Russia is concerned about how its global navigation system, GLONASS — the only one to incorporate leap seconds — would cope, says Vincent Meens of France's National Centre for Space Studies, and the chair of the ITU subgroup tasked with debating the topic. Britain's argument is based largely on the desire to keep a link between official time and Earth's rotation, says Peter Whibberley, a metrologist at the National Physical Laboratory in Teddington, UK.

Astronomers are among those who would be affected if the leap second were to be scrapped. Their software would need to cope with Earth's rotational time — which defines when stars and galaxies are seen in the sky — being offset by more than a second from universal time, says Meens.

On 18 November, the ITU announced that it would defer a decision until 2023 when it will have more information on the impacts of losing the second.

The union did, however, decide to make changes to the international treaty that currently defines UTC, and in turn the leap second. Rather than having a stand-alone definition of UTC, the treaty will cite an SI definition, and mention of the leap second will move to become part of a 'description' of UTC in a subsidiary section of the treaty that expires in 2023.

Whibberley says that the effect will be to remove responsibility for UTC from the ITU, and that the General Conference on Weights and Measures (CGPM) — which is already responsible for defining SI units, including the second — is most likely to become the authority in the future. But the change is unlikely to speed up the decision on whether to scrap the leap second: the CGPM's next chance to even propose a change is not until 2018. ■

Decades of studies on chimpanzee brains and behaviour will be captured in an online resource.

VINCENT J. MUSI/NATL GEOGRAPHIC CREATIVE

BIOMEDICAL RESEARCH

Chimps retire to a digital world

NIH to fund a cache of brain tissue and online data in place of live-animal experimentation.

BY SARA REARDON

Panzee the chimpanzee was a skilled communicator that could tell untrained humans where to find hidden food by using gestures and vocalizations. Austin the chimp was particularly adept with a computer, and scientists have been scanning its genome for clues to its unusual cognitive abilities.

Both apes lived at a language-research centre at Georgia State University in Atlanta, and both died several years ago — but they will live on in an online database of brain scans and behavioural data from nearly 250 chimpanzees.

Researchers hope to combine this trove, now in development, with a biobank of chimpanzee brains to enable scientists anywhere in the world to study the animals' neurobiology.

This push to repurpose old data is especially timely now that the US National Institutes of Health (NIH) has decided to retire its remaining research chimpanzees. The agency decommissioned more than 300 animals in 2013, but kept 50 available for research in case of a public-health emergency. Following an 18 November decision, this remaining population will also be sent to sanctuaries in the coming years. The NIH also hopes to retire another 82 chimps that it supports but does not own, says director Francis Collins.

"We were on a trajectory toward zero, and today's the day we're at zero," says Jeffrey Kahn, a bioethicist at Johns Hopkins University in Baltimore, Maryland, who led a 2011 study on the NIH chimp colony for the Institute of Medicine.

The NIH's latest move, along with a decision in June by the US Fish and Wildlife Service to give research chimps endangered-species protections, effectively ends the possibility of biomedical research on the animals in the United States.

The retirement of the NIH chimps will also end non-invasive studies on the 139 NIH-owned animals at the University of Texas MD Anderson Cancer Center primate facility in Bastrop. Its director, Christian Abee, says that researchers have published more than 50 behavioural studies since 2012 using these animals. "There is no other alternative for cognitive research in chimpanzees," he says.

That makes the NIH-funded chimp database all the more important. "This is a very unique window of opportunity to make sure that there's a legacy and a contribution from the lives they have lived," says project leader Chet Sherwood, a biological anthropologist at George Washington University in Washington DC.

ONLINE LEGACY

In the next few months, Sherwood's team plans to launch a website with a database for researchers and an educational component for the public. The site will eventually include existing data on the chimps' performance in behaviour and personality tests, scans of the primates' brain structure and activity, and their pedigrees and some genetic information. Sherwood and his colleagues plan to model the website on that of the Human Connectome Project, an open-access collection of brain scans from 1,200 individuals that researchers can use to study the links between brain structure and activity and human traits.

The team is also collaborating with the Allen Brain Institute in Seattle, Washington, to create an atlas of gene expression in the chimp brain. Researchers who want to study chimp brains in more detail can request tissue and blood samples from the team, which has nearly 250 preserved organs stored at facilities in Washington DC and Atlanta.

"This is a unique window of opportunity to make sure that there's a legacy and a contribution from the lives they have lived."

But some scientists and advocates worry about the consequences of losing access to research chimps. Frankie Trull, director of the Foundation for Biomedical Research in Washington DC, which advocates for animal research, says that the US government may regret its decision if a public-health threat emerges that would be best studied in chimpanzees. Others caution that the dwindling number of research animals will make it difficult to develop therapies — such as vaccines against Ebola — for wild chimps, which would help both the animals and human beings.

In the meantime, the NIH is struggling to find homes for its newly retired chimps. By law, retired animals are sent to a federal sanctuary known as Chimp Haven in Keithville, Louisiana, but that facility has only 25 places available now. Nearly 310 NIH-owned animals need to be resettled, and Collins says that the agency is still evaluating its options — a situation that worries lawmakers.

On 20 November, two members of Congress sent the NIH a letter asking the agency for its plan to rehome the remaining chimps. "We want to make sure that for the sake of taxpayers and these much-abused chimpanzees, these delays are overcome immediately," they wrote.

Although retired, the apes of Chimp Haven may one day re-enter research labs — posthumously. Sherwood's team is drafting an agreement with the sanctuary to obtain the animals' brains when they die; it also hopes to acquire organs from chimps in zoos and research facilities. "You can imagine 20 years from now, this ageing population won't be here," he says. "If we weren't making the efforts today, there wouldn't be a way to study neurobiology in chimpanzees." ■

ILLUSTRATION BY DAVID PARKINS

ALL TOGETHER NOW

After 25 years of negotiations, all countries are finally set to take steps to limit global warming. A special issue examines the path to the Paris climate summit, and the road beyond.

When more than 190 nations gather in Paris on 30 November to broker an agreement to mitigate climate change, it will be a turning point for the planet.

A successor to the 1997 Kyoto Protocol has been a long time coming. A previous attempt to shape a global agreement fell apart in 2009, at talks in Copenhagen. Now the world is ready to try again, and for the first time, all countries are poised to take action (see page 418). But the history here is sobering: the quest to build a global climate treaty has hit many obstacles over the past 25 years. Its dramatic story is chronicled in a comic starting on page 427.

Although the United Nations aims to limit global warming to 2 °C, a News Feature on page 436 reveals that this will be much harder than many studies have indicated. Part of the difficulty will be ensuring that any treaty leads to actions with lasting global momentum, say climate-policy experts David Victor and James Leape (see page 439). But Johan Rockström, director of the Stockholm Resilience Centre, argues on page 411 that Paris will be a success if it shows that the world is serious about addressing the climate problem.

To explore the backstory to the talks, historian Adam Rome reviews seminal books on sustainability from the 1960s and 1970s (see page 443). A News story explores the challenges facing the Green Climate Fund, a UN mechanism to help developing nations adapt to climate change (see page 419).

Online, *Nature* presents videos about the climate summit, as well as other unique material, at www.nature.com/parisclimate. We will also cover the talks as they happen.

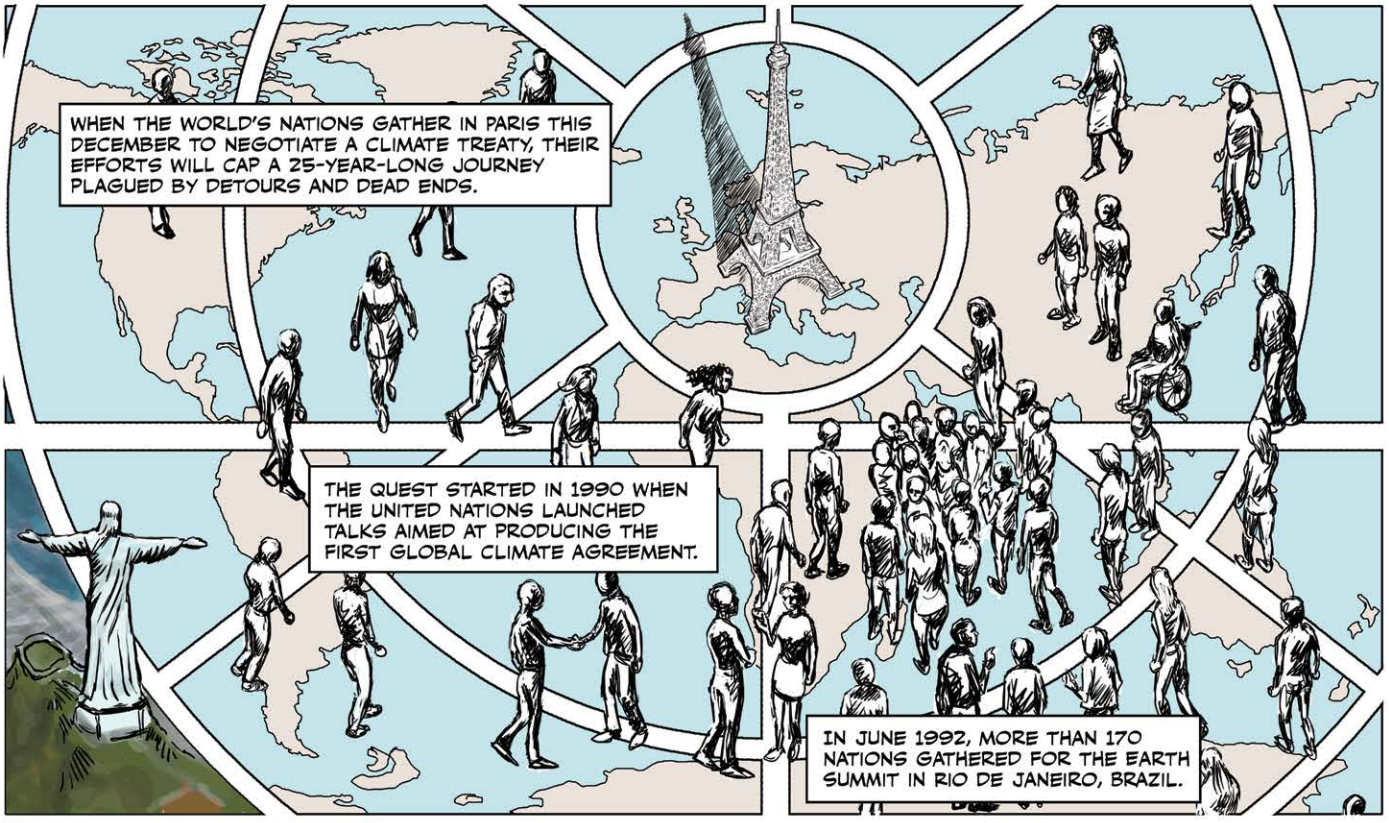
Any agreement reached in Paris will not solve the climate problem, but it could lay a solid foundation for collective action (see page 409). The quest to save the planet will continue for many more years. ■

THE FRAGILE FRAMEWORK

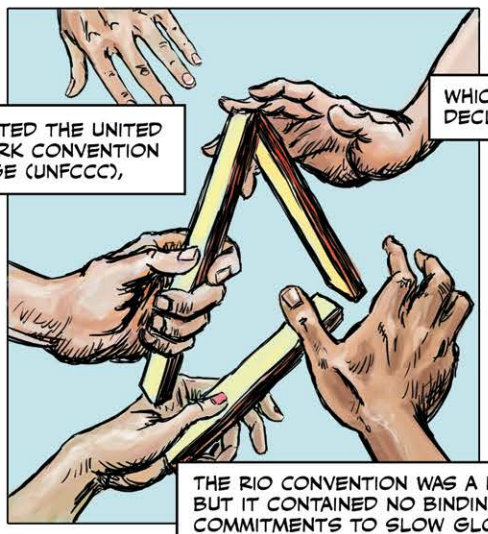
Can nations unite to save Earth's climate?

A COMIC BY
RICHARD MONASTERSKY
AND *NICK SOUSANIS*

PARIS CLIMATE TALKS
A *Nature* special issue
nature.com/parisclimate



IN RIO, THEY ADOPTED THE UNITED NATIONS FRAMEWORK CONVENTION ON CLIMATE CHANGE (UNFCCC),



WHICH DECLARED:

empêche toute perturbation anthropique dangereuse du système climatique

...上气候系统受到危险的人为干扰

The ultimate objective of this Convention … is to achieve … stabilization of greenhouse gas concentrations in the atmosphere at a level that would prevent dangerous anthropogenic interference with the climate system.

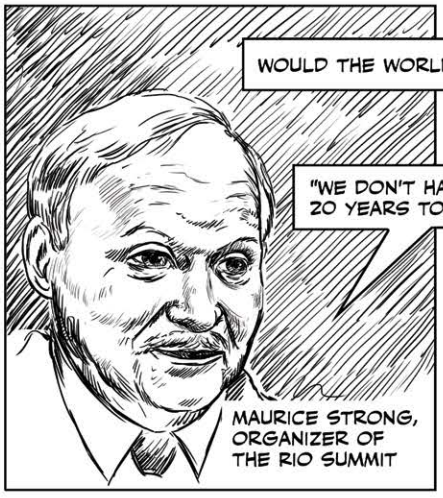
не допускал бы опасного антропогенного воздействия на

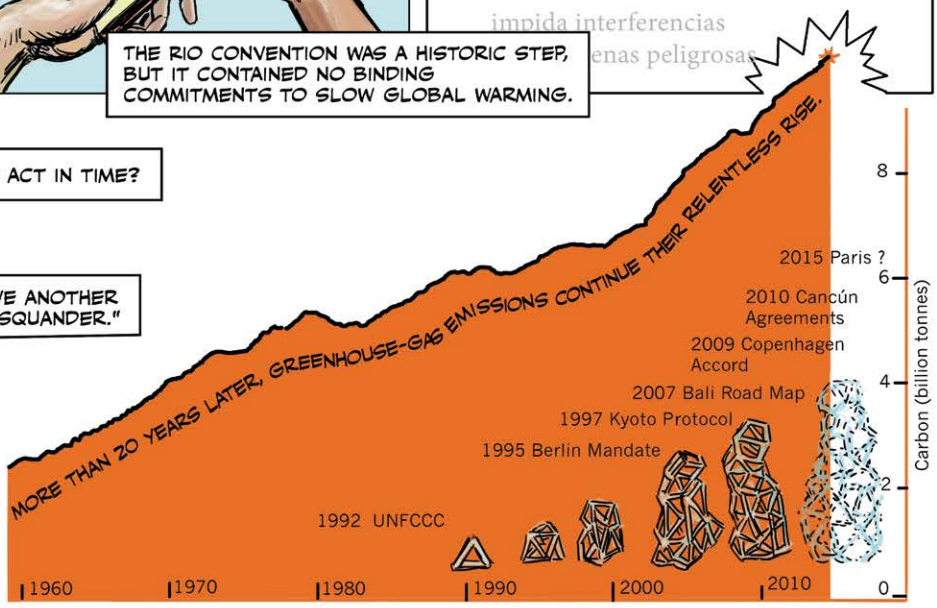
الانسان جانب من خطير تدخل دون يحول

impida interferencias

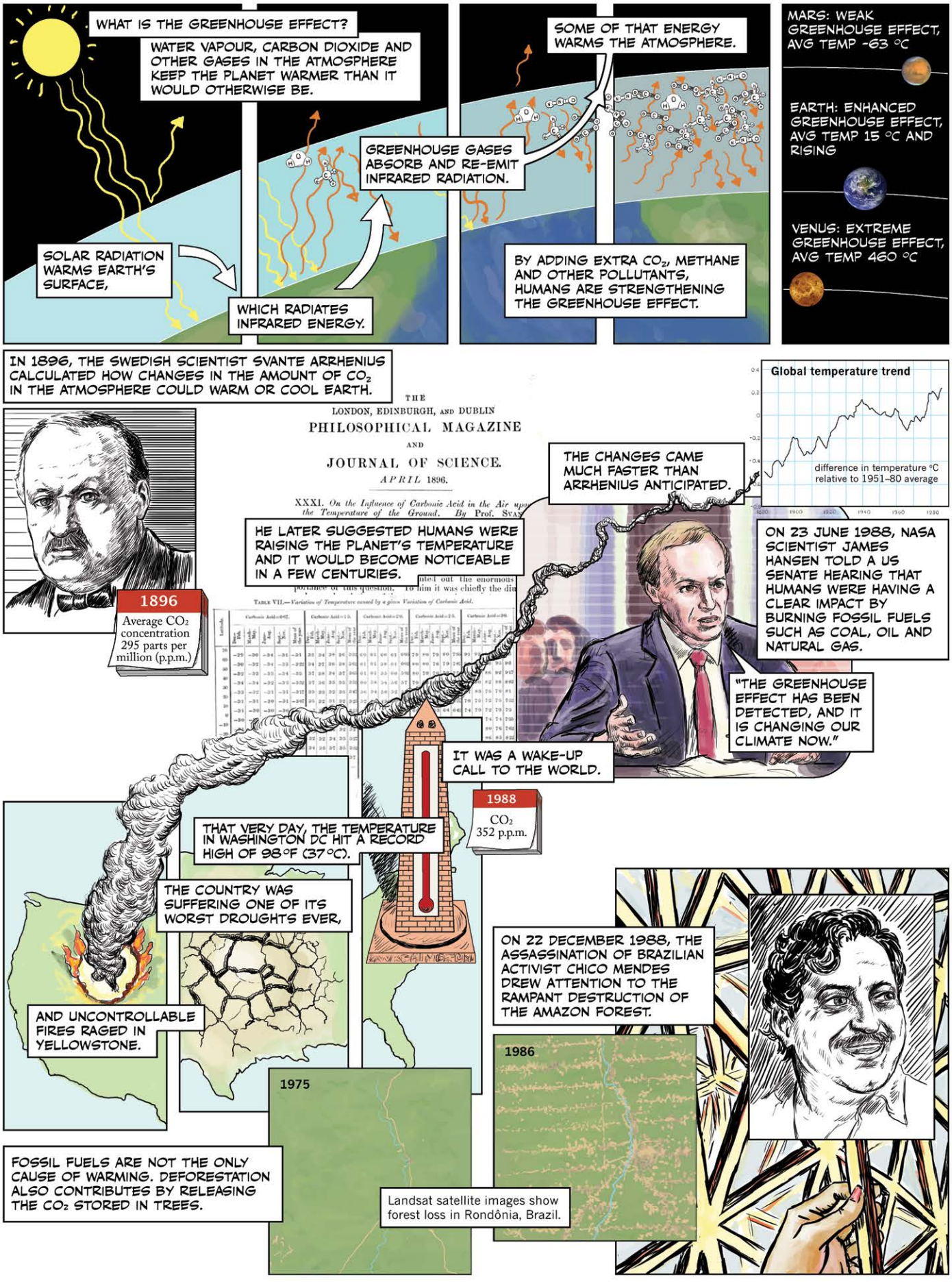
...enas peligrosas

THE RIO CONVENTION WAS A HISTORIC STEP, BUT IT CONTAINED NO BINDING COMMITMENTS TO SLOW GLOBAL WARMING.





FOR SOURCES, CREDITS AND FURTHER READING, SEE: GO.NATURE.COM/EXUQOT


WHAT IS THE GREENHOUSE EFFECT?
WATER VAPOUR, CARBON DIOXIDE AND OTHER GASES IN THE ATMOSPHERE KEEP THE PLANET WARMER THAN IT WOULD OTHERWISE BE.
SOLAR RADIATION WARMS EARTH'S SURFACE,
WHICH RADIATES INFRARED ENERGY.
GREENHOUSE GASES ABSORB AND RE-EMIT INFRARED RADIATION.
SOME OF THAT ENERGY WARMS THE ATMOSPHERE.
BY ADDING EXTRA CO_2, METHANE AND OTHER POLLUTANTS, HUMANS ARE STRENGTHENING THE GREENHOUSE EFFECT.
MARS: WEAK GREENHOUSE EFFECT, AVG TEMP -63 °C
EARTH: ENHANCED GREENHOUSE EFFECT, AVG TEMP 15 °C AND RISING
VENUS: EXTREME GREENHOUSE EFFECT, AVG TEMP 460 °C
IN 1896, THE SWEDISH SCIENTIST SVANTE ARRHENIUS CALCULATED HOW CHANGES IN THE AMOUNT OF CO_2 IN THE ATMOSPHERE COULD WARM OR COOL EARTH.
THE LONDON, EDINBURGH, AND DUBLIN PHILOSOPHICAL MAGAZINE AND JOURNAL OF SCIENCE.
APRIL 1896.
XXXI. On the Influence of Carbonic Acid in the Air upon the Temperature of the Ground. By Prof. SVA
1896
Average CO_2 concentration 295 parts per million (p.p.m.)
HE LATER SUGGESTED HUMANS WERE RAISING THE PLANET'S TEMPERATURE AND IT WOULD BECOME NOTICEABLE IN A FEW CENTURIES.
Global temperature trend
difference in temperature °C relative to 1951–80 average
THE CHANGES CAME MUCH FASTER THAN ARRHENIUS ANTICIPATED.
ON 23 JUNE 1988, NASA SCIENTIST JAMES HANSEN TOLD A US SENATE HEARING THAT HUMANS WERE HAVING A CLEAR IMPACT BY BURNING FOSSIL FUELS SUCH AS COAL, OIL AND NATURAL GAS.
"THE GREENHOUSE EFFECT HAS BEEN DETECTED, AND IT IS CHANGING OUR CLIMATE NOW."
IT WAS A WAKE-UP CALL TO THE WORLD.
1988
CO_2 352 p.p.m.
THAT VERY DAY, THE TEMPERATURE IN WASHINGTON DC HIT A RECORD HIGH OF 98 °F (37 °C).
THE COUNTRY WAS SUFFERING ONE OF ITS WORST DROUGHTS EVER,
AND UNCONTROLLABLE FIRES RAGED IN YELLOWSTONE.
ON 22 DECEMBER 1988, THE ASSASSINATION OF BRAZILIAN ACTIVIST CHICO MENDES DREW ATTENTION TO THE RAMPANT DESTRUCTION OF THE AMAZON FOREST.
1975
1986
Landsat satellite images show forest loss in Rondônia, Brazil.
FOSSIL FUELS ARE NOT THE ONLY CAUSE OF WARMING. DEFORESTATION ALSO CONTRIBUTES BY RELEASING THE CO_2 STORED IN TREES.

ALARMED BY THE GROWING PROBLEM, THE UNITED NATIONS CREATED THE INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE (IPCC) IN 1988 TO ASSESS THE ISSUE.

ipcc

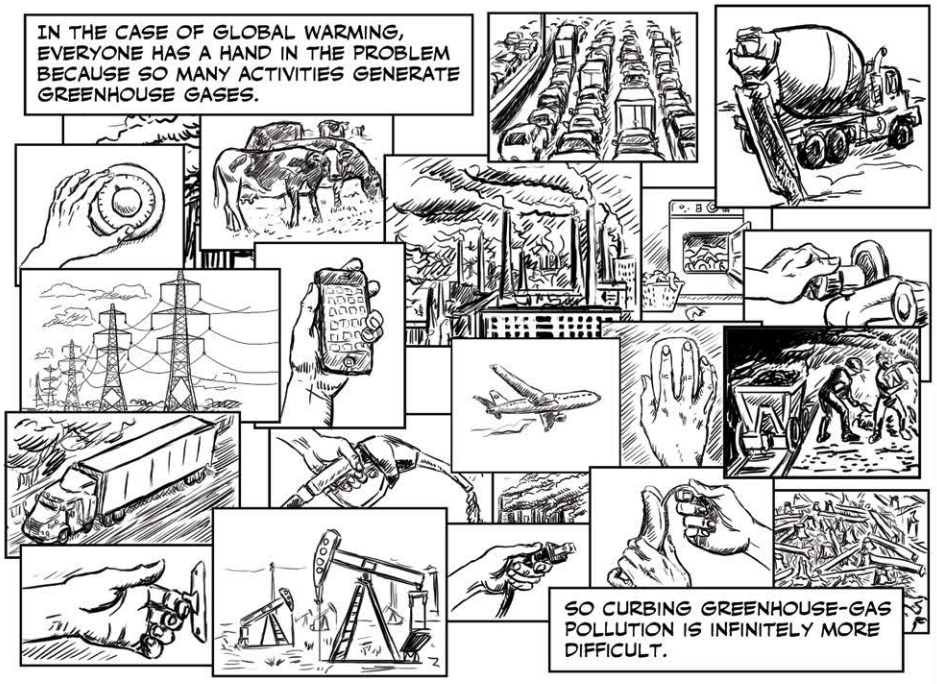
AT THE IPCC'S FIRST MEETING, THE DIRECTOR OF THE UNITED NATIONS ENVIRONMENT PROGRAMME, MOSTAFA TOLBA, IMPLORED SCIENTISTS TO USE THE TIME LEFT IN THE CENTURY – JUST 4,000 DAYS – TO DEAL WITH CLIMATE CHANGE.

HOPES WERE HIGH BECAUSE THE WORLD HAD ALREADY TAKEN STEPS TO SOLVE ONE ENVIRONMENTAL DISASTER.

IN 1987, NATIONS ADOPTED A TREATY TO PROTECT THE OZONE LAYER.

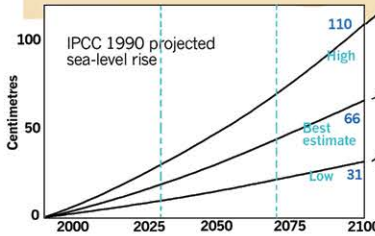
BUT REACHING THAT AGREEMENT WAS RELATIVELY EASY BECAUSE ONLY A HANDFUL OF COMPANIES IN A FEW COUNTRIES PRODUCED OZONE-DESTROYING COMPOUNDS.

IN THE CASE OF GLOBAL WARMING, EVERYONE HAS A HAND IN THE PROBLEM BECAUSE SO MANY ACTIVITIES GENERATE GREENHOUSE GASES.



SO CURBING GREENHOUSE-GAS POLLUTION IS INFINITELY MORE DIFFICULT.

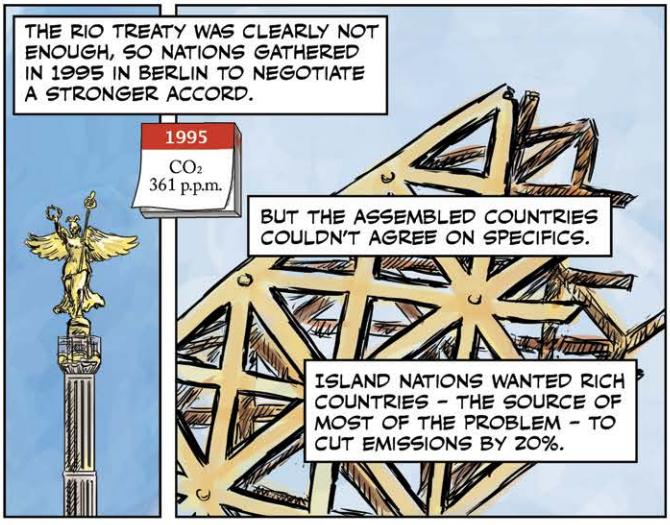
IN ITS FIRST REPORT, THE IPCC FORECASTED THAT IF CURRENT TRENDS CONTINUE UNTIL 2100, THE WORLD WOULD BE 4 °C WARMER THAN IT WAS IN 1850. SWELLING OCEANS WOULD BE A MAJOR PROBLEM ...



...BECAUSE HALF OF HUMANITY INHABITS COASTAL REGIONS.

A MONSTROUS CYCLONE DROVE THAT POINT HOME IN 1991 WHEN IT KILLED MORE THAN 140,000 PEOPLE IN BANGLADESH.

THE RIO TREATY WAS CLEARLY NOT ENOUGH, SO NATIONS GATHERED IN 1995 IN BERLIN TO NEGOTIATE A STRONGER ACCORD.



BUT THE ASSEMBLED COUNTRIES COULDN'T AGREE ON SPECIFICS.

ISLAND NATIONS WANTED RICH COUNTRIES – THE SOURCE OF MOST OF THE PROBLEM – TO CUT EMISSIONS BY 20%.

THE US DEMANDED THAT DEVELOPING NATIONS AGREE TO FUTURE COMMITMENTS.

CHINA SAID DEVELOPED NATIONS HAD TO ACT FIRST BECAUSE THEY HAD CAUSED THE PROBLEM.

DURING ALL-NIGHT NEGOTIATIONS, GERMANY'S ENVIRONMENT MINISTER, ANGELA MERKEL, BROKERED A DEAL. COUNTRIES WOULD HAVE TWO YEARS TO AGREE ON EMISSIONS LIMITS FOR DEVELOPED NATIONS.

IN DECEMBER 1997, COUNTRIES GATHERED IN KYOTO, JAPAN, TO HASH OUT A NEW TREATY. BUT THEY COULDN'T AGREE ON HOW MUCH DEVELOPED NATIONS SHOULD TRIM THEIR EMISSIONS.

EUROPEAN UNION CALLED FOR A 15% CUT

ISLAND NATIONS DEMANDED A 20% CUT

JAPAN PROPOSED A 5% CUT

THE US WANTED DEVELOPING COUNTRIES TO ACT, TOO.

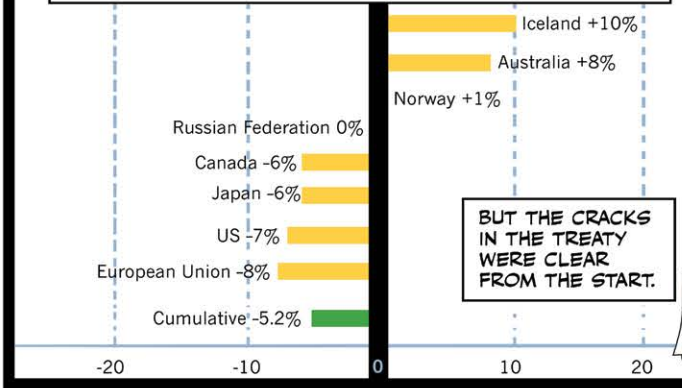
AFTER WORKING THROUGH THE FINAL NIGHT, NEGOTIATORS REACHED AN AGREEMENT CALLED THE KYOTO PROTOCOL. IT WAS THE FIRST TIME THAT COUNTRIES PROMISED TO REIN IN GREENHOUSE-GAS POLLUTION BY SPECIFIC AMOUNTS.

THE KYOTO PROTOCOL SPLIT THE WORLD IN TWO: INDUSTRIALIZED COUNTRIES WITH EMISSIONS LIMITS ...

... AND DEVELOPING COUNTRIES WITHOUT.

THE PROTOCOL ALSO ALLOWED FOR FLEXIBILITY IN HOW COUNTRIES MET THEIR COMMITMENTS. DEVELOPED NATIONS COULD GET CREDIT FOR REDUCING EMISSIONS IN POORER ONES.

DEVELOPED COUNTRIES PROMISED TO CUT THEIR OVERALL EMISSIONS TO 5.2% BELOW 1990 LEVELS FOR THE PERIOD 2008–12. EACH COUNTRY HAD ITS OWN TARGET.



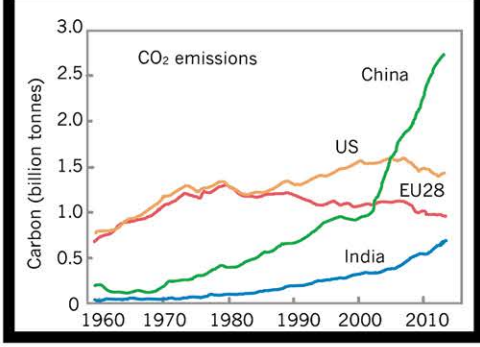
BUT THE CRACKS IN THE TREATY WERE CLEAR FROM THE START.

1998
CO_2 367 p.p.m.

THE US REFUSED TO RATIFY THE PACT BECAUSE OF CONCERNS THAT ITS ECONOMY WOULD SUFFER WHILE DEVELOPING NATIONS INCREASED THEIR POLLUTION WITHOUT LIMITS.

IN 2001, US PRESIDENT GEORGE W. BUSH REJECTED THE AGREEMENT, SAYING "THE KYOTO PROTOCOL WAS FATALLY FLAWED IN FUNDAMENTAL WAYS."

SOON, WORLD EVENTS MADE CLEAR HOW LIMITED THE PROTOCOL WAS. IN 2006, CHINA PASSED THE US TO BECOME THE WORLD'S LARGEST CARBON EMITTER.



CANADA FORMALLY WITHDREW FROM KYOTO IN 2011.

Planet Daily
2014 HOTTEST YEAR ON RECORD

Planet Daily
2010 HOTTEST YEAR ON RECORD

Planet Daily
2005 HOTTEST YEAR ON RECORD

Planet Daily
1998 HOTTEST YE

WASHINGTON – Earth was warmer in 1998 than in any other year on record, climate scientists reported Thursday. Studies of air and ocean temperatures by researchers in England and the Unit[illegible]tes both show that the average

Th[illegible] average for the rest of the century, capping a general warming trend over the past few decades. Researchers blame the rising temperatures on greenhouse gasses

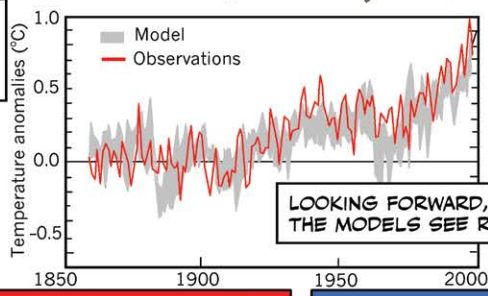
MEANWHILE, GLOBAL TEMPERATURES CONTINUED TO SOAR.

THROUGHOUT THE CLIMATE NEGOTIATIONS, SCIENTISTS HAVE TRIED TO SHOW WHAT KIND OF WORLD AWAITS FUTURE GENERATIONS IF GLOBAL WARMING CONTINUES.

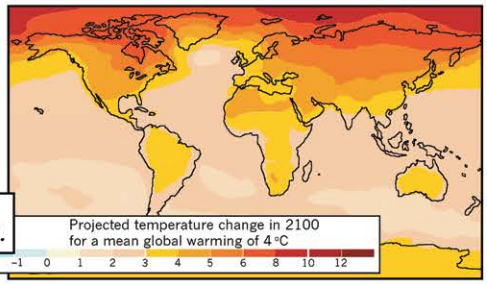
SUCH FORECASTS COME FROM COMPLEX CLIMATE SYSTEM MODELS, WHICH DIVIDE THE GLOBE INTO MILLIONS OF CELLS ...

... AND SIMULATE THE ATMOSPHERE, OCEANS AND BIOSPHERE IN 3D.

RESEARCHERS HAVE CONFIDENCE IN THEIR MODELS BECAUSE THEY CAN REPRODUCE FEATURES OF PAST AND CURRENT CLIMATES.



LOOKING FORWARD, THE MODELS SEE RED.



ALTHOUGH SCIENTISTS AGREED HUMANS WERE WARMING THE PLANET, SOME POLITICIANS DENIED THAT FACT.

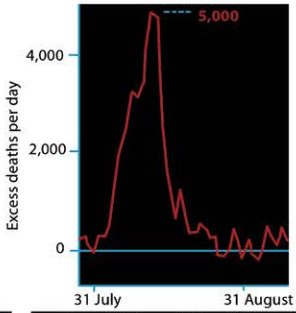
"WITH ALL OF THE HYSTERIA, ALL OF THE FEAR, ALL OF THE PHONY SCIENCE, COULD IT BE THAT MAN-MADE GLOBAL WARMING IS THE GREATEST HOAX EVER PERPETRATED ON THE AMERICAN PEOPLE?" – US SENATOR JAMES INHOFE, 2003

OTHER POLITICIANS DEFENDED CLIMATE SCIENCE.

"WE SHOULD INVESTIGATE THE WELL-FUNDED EFFORT BY CERTAIN OIL COMPANIES TO MANUFACTURE CONTROVERSY AND CAST DOUBT ON THE REALITY OF GLOBAL WARMING." – US REPRESENTATIVE HENRY WAXMAN, 2006

2003
CO_2
376 p.p.m.

IN 2003, EUROPE SUFFERED A PROLONGED HEAT WAVE THAT KILLED AN ESTIMATED 70,000 PEOPLE.



THE IMPACTS WERE GETTING CLEARER. THE IPCC DECLARED IN 2007: "WARMING OF THE CLIMATE SYSTEM IS UNEQUIVOCAL."

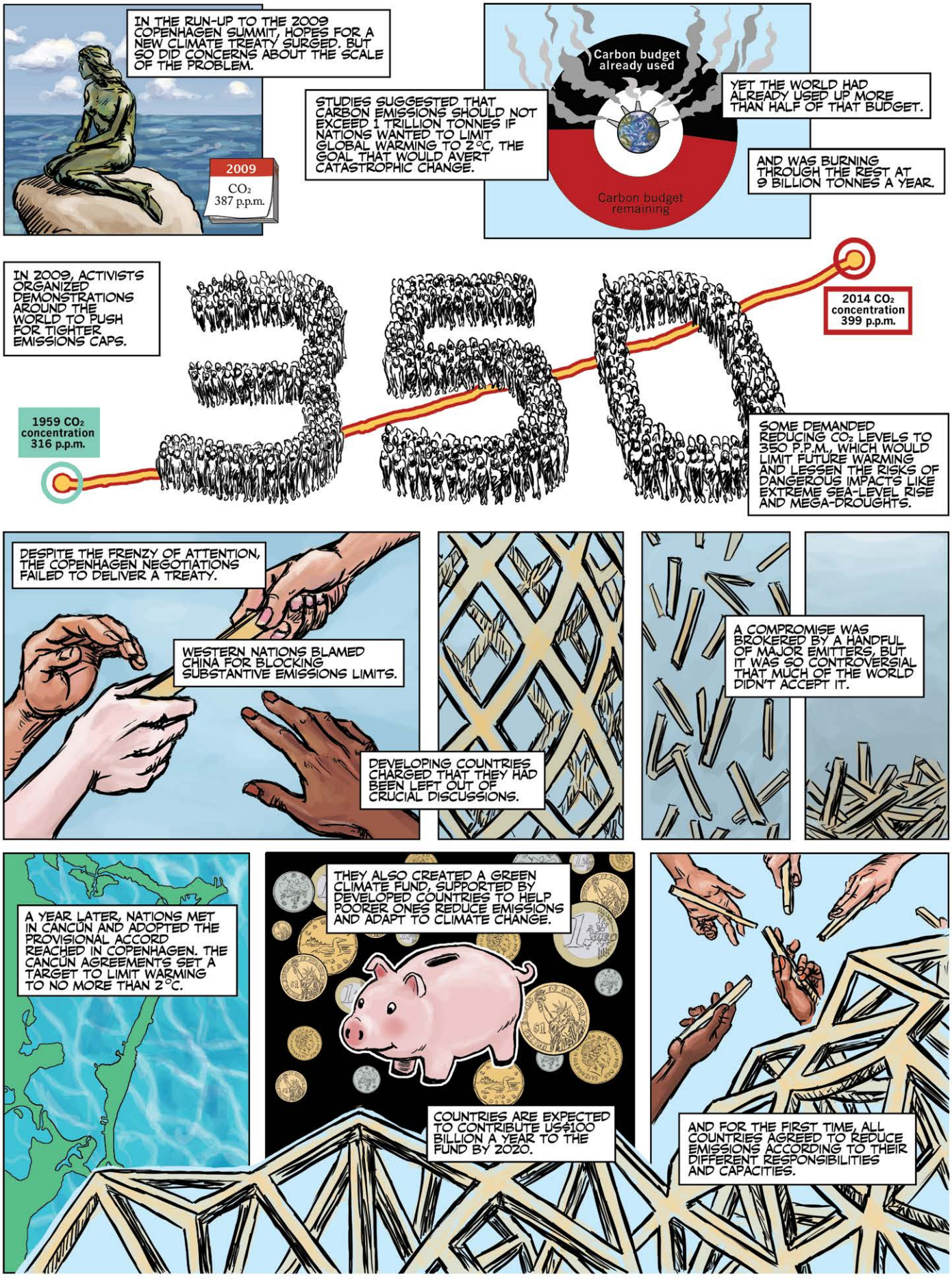
LATER THAT YEAR, THE IPCC WAS AWARDED THE NOBEL PEACE PRIZE FOR ITS EFFORTS.

WHILE THE SCIENCE RACED AHEAD, THE NEGOTIATIONS DRAGGED ON. NATIONS MET IN BALI IN DECEMBER 2007.

FOR THE FIRST TIME, DEVELOPING NATIONS AGREED TO "MITIGATION ACTIONS" OF THEIR OWN CHOOSING TO LIMIT CLIMATE CHANGE.

THE TALKS WERE SO FRACTIOUS THAT AT ONE POINT, THE CHAIRMAN BROKE DOWN IN TEARS. BUT A DEADLINE WAS SET FOR A TREATY IN 2009 THAT WOULD INCLUDE NEW COMMITMENTS BY DEVELOPED COUNTRIES.

AND PARTICIPANTS DECIDED THE NEXT TREATY WOULD ADDRESS DEFORESTATION.


IN THE RUN-UP TO THE 2009 COPENHAGEN SUMMIT, HOPES FOR A NEW CLIMATE TREATY SURGED. BUT SO DID CONCERNS ABOUT THE SCALE OF THE PROBLEM.
2009
CO_2 387 p.p.m.
STUDIES SUGGESTED THAT CARBON EMISSIONS SHOULD NOT EXCEED 1 TRILLION TONNES IF NATIONS WANTED TO LIMIT GLOBAL WARMING TO 2 °C, THE GOAL THAT WOULD AVERT CATASTROPHIC CHANGE.
Carbon budget already used
Carbon budget remaining
YET THE WORLD HAD ALREADY USED UP MORE THAN HALF OF THAT BUDGET.
AND WAS BURNING THROUGH THE REST AT 9 BILLION TONNES A YEAR.
IN 2009, ACTIVISTS ORGANIZED DEMONSTRATIONS AROUND THE WORLD TO PUSH FOR TIGHTER EMISSIONS CAPS.
1959 CO_2 concentration 316 p.p.m.
2014 CO_2 concentration 399 p.p.m.
SOME DEMANDED REDUCING CO_2 LEVELS TO 350 P.P.M., WHICH WOULD LIMIT FUTURE WARMING AND LESSEN THE RISKS OF DANGEROUS IMPACTS LIKE EXTREME SEA-LEVEL RISE AND MEGA-DROUGHTS.
DESPITE THE FRENZY OF ATTENTION, THE COPENHAGEN NEGOTIATIONS FAILED TO DELIVER A TREATY.
WESTERN NATIONS BLAMED CHINA FOR BLOCKING SUBSTANTIVE EMISSIONS LIMITS.
DEVELOPING COUNTRIES CHARGED THAT THEY HAD BEEN LEFT OUT OF CRUCIAL DISCUSSIONS.
A COMPROMISE WAS BROKERED BY A HANDFUL OF MAJOR EMITTERS, BUT IT WAS SO CONTROVERSIAL THAT MUCH OF THE WORLD DIDN'T ACCEPT IT.
A YEAR LATER, NATIONS MET IN CANCÚN AND ADOPTED THE PROVISIONAL ACCORD REACHED IN COPENHAGEN. THE CANCÚN AGREEMENTS SET A TARGET TO LIMIT WARMING TO NO MORE THAN 2 °C.
THEY ALSO CREATED A GREEN CLIMATE FUND, SUPPORTED BY DEVELOPED COUNTRIES TO HELP POORER ONES REDUCE EMISSIONS AND ADAPT TO CLIMATE CHANGE.
COUNTRIES ARE EXPECTED TO CONTRIBUTE US$100 BILLION A YEAR TO THE FUND BY 2020.
AND FOR THE FIRST TIME, ALL COUNTRIES AGREED TO REDUCE EMISSIONS ACCORDING TO THEIR DIFFERENT RESPONSIBILITIES AND CAPACITIES.

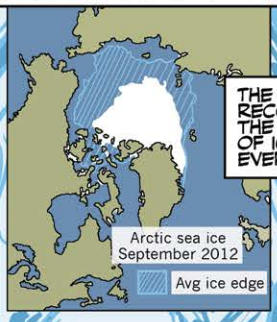
AS THE NEGOTIATIONS HAVE CRAWLED ALONG, THE WORLD HAS HURTLED THROUGH CHANGES.

IN OCTOBER 2011, THE GLOBAL POPULATION TOPPED 7 BILLION PEOPLE.

A HEAT WAVE IN RUSSIA IN 2010 KILLED ROUGHLY 55,000 PEOPLE AND STOKED WILDFIRES ACROSS THE COUNTRY.

IN 2012, 97% OF THE GREENLAND ICE SHEET SHOWED SIGNS OF MELTING – THE FIRST TIME SUCH AN EXTENSIVE AREA HAD THAWED.

NOVEMBER 2013: SUPERSTORM HAIYAN BECAME THE STRONGEST ON RECORD TO MAKE LANDFALL WHEN IT SLAMMED INTO THE PHILIPPINES WITH WINDS OF 315 KILOMETRES PER HOUR.



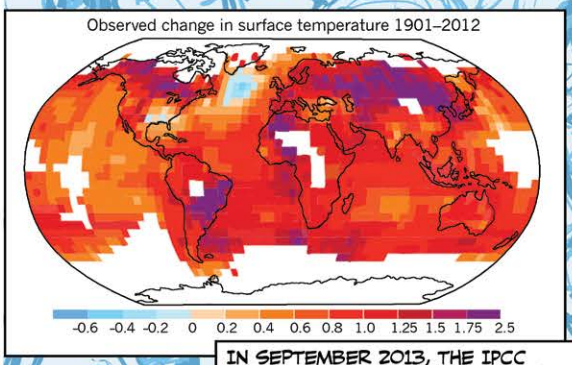
THE ARCTIC SEA SET A RECORD IN 2012 FOR THE LEAST AMOUNT OF ICE COVERAGE EVER DOCUMENTED.

AT CLIMATE TALKS IN 2013, PHILIPPINE DELEGATE NADEREV SANO BROKE DOWN OVER HIS COUNTRY'S DEVASTATION AND THE LACK OF PROGRESS IN NEGOTIATIONS.

"BY FAILING TO MEET THE OBJECTIVE OF THE CONVENTION, WE MAY HAVE RATIFIED THE DOOM OF VULNERABLE COUNTRIES."

9 MAY 2013: DAILY ATMOSPHERIC MEASUREMENTS OF CO_2 AT A STATION IN HAWAII TOPPED 400 P.P.M. FOR THE FIRST TIME IN, PERHAPS, MILLIONS OF YEARS.

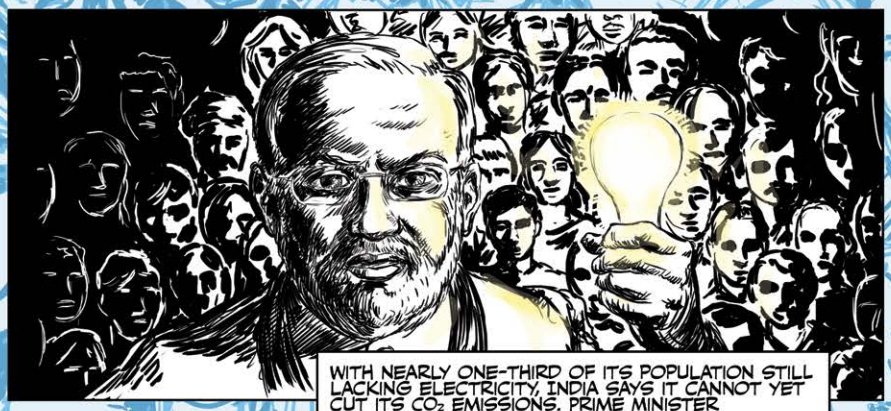
JAMES HANSEN, WHO HAD SOUNDED THE ALARM OVER GLOBAL WARMING, WAS ARRESTED SEVERAL TIMES FOR PROTESTS OVER THE PROPOSED KEYSTONE XL PIPELINE.



IN SEPTEMBER 2013, THE IPCC REPORTED THAT "HUMAN INFLUENCE ON THE CLIMATE SYSTEM IS CLEAR."

2014: CHINA OVERTOOK THE EUROPEAN UNION IN PER CAPITA EMISSIONS OF CO_2.

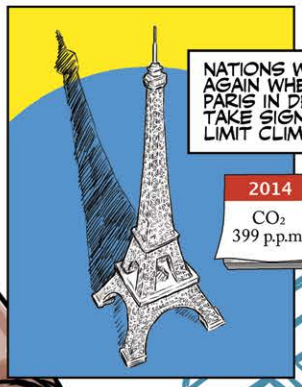
THE US AND CHINA BROKERED A HISTORIC CLIMATE DEAL IN NOVEMBER 2014. THE US PLEDGED TO CUT GREENHOUSE-GAS EMISSIONS 26–28% BELOW 2005 LEVELS BY 2025. CHINA SAID ITS EMISSIONS WOULD PEAK AROUND 2030.



THESE BECAME THE BASIS FOR WHAT THE TWO COUNTRIES SUBMITTED TO THE UNFCCC AS THEIR PROPOSED EMISSIONS TARGETS IN ADVANCE OF THE PARIS SUMMIT. SO FAR, MORE THAN 160 COUNTRIES HAVE SUBMITTED THEIR PLEDGES.

WITH NEARLY ONE-THIRD OF ITS POPULATION STILL LACKING ELECTRICITY, INDIA SAYS IT CANNOT YET CUT ITS CO_2 EMISSIONS. PRIME MINISTER NARENDRA MODI'S GOVERNMENT PLEDGED TO SUBSTANTIALLY INCREASE ENERGY EFFICIENCY.

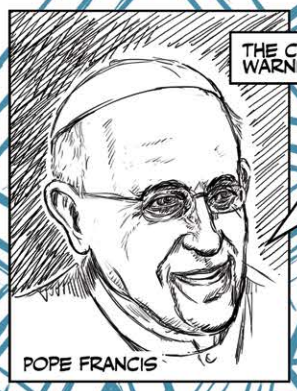
2014 SET THE RECORD AS THE HOTTEST YEAR EVER RECORDED FOR GLOBAL LAND AND SEA SURFACE TEMPERATURES.



NATIONS WILL FACE A TEST AGAIN WHEN THEY MEET IN PARIS IN DECEMBER: CAN THEY TAKE SIGNIFICANT STEPS TO LIMIT CLIMATE CHANGE?

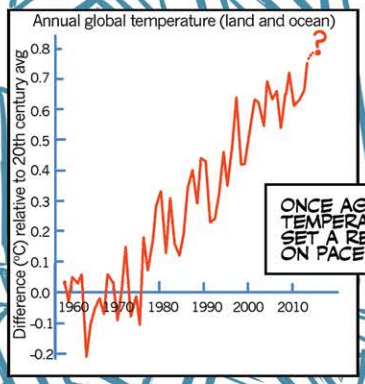
2014
CO_2
399 p.p.m.

THE CALLS FOR ACTION – AND THE WARNING SIGNS – GROW STRONGER.



"CLIMATE CHANGE IS A GLOBAL PROBLEM WITH GRAVE IMPLICATIONS .... IT REPRESENTS ONE OF THE PRINCIPAL CHALLENGES FACING HUMANITY IN OUR DAY."

POPE FRANCIS



ONCE AGAIN, GLOBAL TEMPERATURES WILL SET A RECORD. 2015 IS ON PACE TO TOP 2014.

"ALTHOUGH WE'RE MOVING IN THE RIGHT DIRECTION, IT IS CLEARLY NOT ENOUGH."

CHRISTIANA FIGUERES, EXECUTIVE SECRETARY OF THE UNFCCC

FIRES IN INDONESIA RAVAGED THE COUNTRY AND PUMPED HALF A BILLION TONNES OF CARBON INTO THE AIR, MORE THAN JAPAN PRODUCES IN A YEAR.

THE PARIS PLEDGES WILL PROBABLY LIMIT WARMING TO BELOW 3°C. BUT MUCH STRONGER ACTION IS NEEDED TO STAY UNDER 2°C. COUNTRIES WILL PROBABLY BLOW THROUGH THE TRILLION-TONNE CARBON BUDGET BEFORE 2040.

CO_2 CONTINUES TO THICKEN THE SKIES. CONCENTRATIONS MAY NEVER AGAIN DROP BELOW 400 P.P.M.

EVEN IF GLOBAL WARMING DOES NOT PASS 2°C, THE WORLD MIGHT STILL FACE CALAMITOUS IMPACTS, LIKE PARTS OF THE ANTARCTIC ICE SHEET SLIDING INTO THE OCEAN WITHIN A FEW CENTURIES.

THE WORLD HAS COME A LONG WAY SINCE KYOTO. THIS YEAR, ALL COUNTRIES ARE SETTING THEIR OWN GOALS, WHICH MAY MAKE THE TARGETS MORE REALISTIC. AND NATIONS MAY AGREE TO STEP UP THEIR COMMITMENTS PERIODICALLY.

THE FRAMEWORK TO FIX THE PLANET IS COMING TOGETHER, BUT IT IS FRAGILE AND FAR TOO SMALL. THE JOB OF FINISHING THE TASK WILL FALL TO FUTURE GENERATIONS.

THE FRAGILE FRAMEWORK

ORIGINALLY PUBLISHED IN
NATURE **527**, 427–435 (2015)

For more on the Paris climate talks, see: nature.com/parisclimate

The 2 °C dream

Countries have pledged to limit global warming to 2 °C, and climate models say that is still possible. But only with heroic — and unlikely — efforts.

BY JEFF TOLLEFSON

NIK SPENCER/*NATURE*

The year is 2100 and the world looks nothing like it did when global leaders gathered for the historic climate summit in Paris at the end of 2015. Nearly 8.8 billion people now crowd the planet. Energy consumption has nearly doubled, and economic production has increased more than sevenfold. Vast disparities in wealth remain, but governments have achieved one crucial goal: limiting global warming to 2 °C above pre-industrial temperatures.

The United Nations meeting in Paris proved to be a turning point. After forging a climate treaty, governments immediately moved to halt tropical deforestation and to expand forests around the globe. By 2020, plants and soils were stockpiling more than 17 billion tonnes of extra carbon dioxide each year, offsetting 50% of global CO_2 emissions. Several million wind turbines were installed, and thousands of nuclear power plants were built. The solar industry ballooned, overtaking coal as a source of energy in the waning years of the twenty-first century.

But it took more than this. Governments had to drive emissions into negative territory — essentially sucking greenhouse gases from the skies — by vastly increasing the use of bioenergy, capturing the CO_2 generated and then pumping it underground on truly massive scales. These efforts pulled Earth back from the brink. Atmospheric CO_2 concentrations peaked in 2060, below the target of 450 parts per million (p.p.m.) and continue to fall.

That scenario for conquering global warming is one possible — if optimistic — vision of the future. It was developed by modellers at the Joint Global Change Research Institute in College Park, Maryland, as part of a broad effort by climate scientists to chart possible paths for limiting global warming to 2 °C, a target enshrined in the UN climate convention that will produce the Paris treaty.

Climate modellers have developed dozens of rosy 2 °C scenarios over several years, and these fed into the latest assessment by the Intergovernmental Panel on Climate Change (IPCC). The panel seeks to be policy-neutral and has never formally endorsed the 2-degree target, but its official message, delivered in April 2014, was clear: the goal is ambitious but achievable.

This work has fuelled hope among policymakers and environmentalists, and it will provide a foundation for debate as governments negotiate a new climate agreement at the UN's 2015 Paris Climate Conference starting on 30 November. Despite broad agreement that the emissions-reduction commitments that countries have offered up so far are insufficient, policymakers continue to talk about bending the emissions curve downwards to remain on the path to 2 degrees that was laid out by the IPCC.

But take a closer look, some scientists argue, and the 2 °C scenarios that define that path seem so optimistic and detached from current political realities that they verge on the farcical. Although the caveats and uncertainties are all spelled out in the scientific literature, there is concern that the 2 °C modelling effort has distorted the political debate by obscuring the scale of the challenge. In particular, some researchers have questioned the viability of large-scale bioenergy use with carbon capture and storage (CCS), on which many models now rely as a relatively cheap way to provide substantial negative emissions. The entire exercise has opened up a rift in the scientific community, with some people raising ethical questions about whether scientists are bending to the will of politicians and government funders who want to maintain 2 °C as a viable political target.

"Nobody dares say it's impossible," says Oliver Geden, head of the European Union Research Division at the German Institute for International and Security Affairs in Berlin. "Everybody is sort of underwriting the 2-degree cheque, but scientists have to think about the credibility of climate science."

Modellers are first to acknowledge the limits of their work, and say that the effort is designed to explore options, not predict the future. "We'll tell you how many nuclear power plants you need, or how much CCS, but we can't tell you whether society is going to be willing to do that or not," says Leon Clarke, a senior scientist and modeller at the Joint Global Change Research Institute. "That's a different question."

ONE TRILLION TONNES

The idea of limiting global warming to 2 °C dates back to 1975, when economist William Nordhaus of Yale University in New Haven, Connecticut, proposed that more than 2 or 3 degrees of warming would push the planet outside the temperature range of the past several hundred thousand years. In 1996, the EU adopted that limit, and the Group of 8 (G8) nations signed on in 2009. The parties to the UN convention on climate change affirmed the target in 2009 at their Copenhagen summit, and then formally adopted it a year later in Cancún, Mexico.

The move caught scientists off guard. Before 2009, most modellers had focused on scenarios in which atmospheric CO_2 concentrations stabilized around 550 p.p.m. — double the pre-industrial level — which would probably limit warming to a little less than 3 °C. But as political interest in the 2 °C target grew, a few started exploring the implications. In April 2009, a team led by Myles Allen, a climate scientist at the University of Oxford, UK, published[1] a study concluding that humans would have to limit their total cumulative carbon emissions to 1 trillion tonnes — more than half of which had already been dumped into the atmosphere — to maintain a chance of limiting warming to 2 °C. This trillion-tonne carbon budget provided a scientific baseline for what was now a politically important target, and many modellers shifted gears.

"There were very few scenarios with stringent targets such as 2 °C, and then sponsors started demanding it," says Massimo Tavoni, deputy coordinator of climate-change programmes at the Eni Enrico Mattei Foundation in Milan, Italy.

The flurry of modelling efforts that followed split into two main camps: pay early or pay late (see 'Two paths to 2 °C'). In the former, nations need to slash greenhouse-gas emissions immediately; in the latter, they can buy time for a slower phase-out by developing a massive infrastructure to suck CO_2 out of the air.

"Models that have these negative emissions really do let you continue to party on now, because you have these options later," says John Reilly, co-director of the Joint Program on the Science and Policy of Global Change at the Massachusetts Institute of Technology (MIT) in Cambridge.

> *"It's just simple arithmetic: the carbon budget is so small that you need to go negative."*

In the pay-later approach, most models rely on a combination of bioenergy and CCS. The system starts with planting crops that are harvested and either processed to make biofuels or burnt to generate electricity, which provide carbon-neutral power because the plants absorb CO_2 as they grow. The CO_2 created when the plants are processed is captured and pumped underground, and the process as a whole eats up more emissions than it creates. A consortium sponsored by the US Department of Energy has tested such a system at one facility that produces bioethanol fuel in Illinois, but neither bioenergy nor CCS has been demonstrated on anywhere near the scales imagined by the models.

"It's just simple arithmetic: the carbon budget is so small that you need to go negative, or at least you need to offset some of your emissions in order to get to zero," says Tavoni. "We tried to be honest, and pretty agnostic about whether these transformations are easily achievable."

On the basis of those models and other information, the IPCC estimates that climate mitigation would reduce the projected global consumption in 2100 by 3–11% — a relatively modest amount that would allow the global economy to keep growing overall. But remove either bioenergy or CCS from the scenarios and the costs increase substantially. If mitigation is delayed or bioenergy and CCS are constrained, most models simply can't limit warming to 2 °C.

The question is whether any of those models accurately reflect technical and social challenges. MIT has a model that tends to project costs two or three times the average reported by the IPCC, in part because it tries to reflect difficulties in scaling up any technology, such as the availability of skilled labour and natural resources in different regions. And then there are the technical hurdles. Capturing CO_2 from power plants has proved more difficult and expensive than many had hoped. Just one commercial project is currently operating, at the Boundary Dam Power Station in Saskatchewan, Canada.

Moreover, Reilly says, the number of models that actually completed 2 °C scenarios remains relatively small, and they probably project lower mitigation costs than those that are not able to generate these low-emissions scenarios. "It's a very self-selecting set of models."

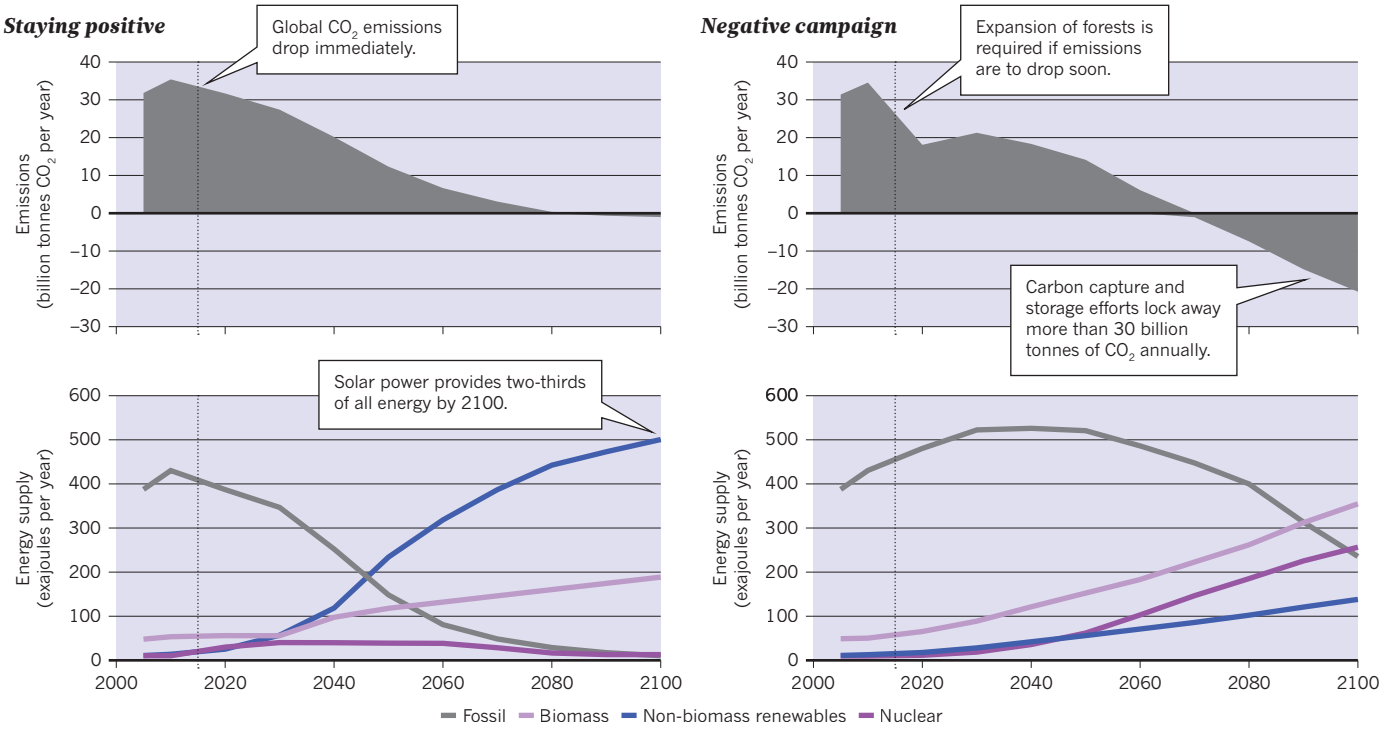
Although the caveats are listed in the IPCC assessment, the report does not adequately highlight economic and technical challenges or modelling uncertainties, says David Victor, a political scientist at the University of California, San Diego, who participated in the IPCC assessment. Victor does not place all the blame on scientists glossing over the problems: when researchers drafted the assessment's chapter on emissions scenarios and costs, he says, they included clear statements about the difficulty of achieving the 2 °C goal. But the governments — led by the EU and a bloc of developing countries — pushed for a more optimistic assessment in the final IPCC report. "We got a lot of pushback, and the text basically got mangled," Victor says.

For all of the concerns and criticisms, however, modellers say that the

Two paths to 2 °C

Modellers have explored various scenarios for limiting global warming to 2 °C. One (left) immediately slashes fossil-fuel use while ramping up renewable-energy use. Another strategy (right) allows continued use of fossil fuels, but bioenergy supplies a growing share of energy. Carbon from the bioenergy industry is captured and stored, driving overall emissions below zero.

SOURCE: IIASA/IPCC



exercises have illuminated important research questions, such as how much bioenergy and CCS will cost and what effects they will have on land use, food systems and water availability.

One 2014 study[2] in *Earth's Future*, for instance, found that it would be difficult to grow enough bioenergy crops, even with second-generation cellulosic biofuels, which are made not only from a plant's sugars but also from the carbon in its stem and woody materials. The effort would require significant boosts in crop yields and the use of 77% more nitrogen fertilizer by 2100. The bioenergy would also need to be produced in centralized facilities that capture the bulk of the emissions. Unless everything goes right, scaling up to the level projected in many models would be difficult without significantly reducing food production or clearing large swathes of natural ecosystems for farmland.

"If we need to ramp up such a large infrastructure, we need to investigate what that implies," says Sabine Fuss, an environmental scientist at the Mercator Research Institute on Global Commons and Climate Change in Berlin.

Fuss led a commentary[3] in *Nature Climate Change* in October 2014 calling for a transdisciplinary research agenda on negative emissions. One of the first outgrowths of that work, led by co-author Peter Smith, a biologist at the University of Aberdeen, UK, is an upcoming assessment of carbon-negative strategies and potential limitations. Strategies include bioenergy with CCS, as well as other ways of absorbing carbon, such as planting forests, using chemical scrubbers to capture CO_2 directly from the air and crushing rocks to enhance geological weathering that consumes the gas.

"The science behind these technologies is probably a bit behind the models," Smith says. "This sort of provides a road map for where we need to go in the next two or three years."

RISK FACTOR

Modellers are also digging into real-world complexities. Most models assume that participation in climate mitigation will be global, that countries will put a common price on carbon, that technological solutions will be widely available and that this combination will drive investment towards relatively cheap mitigation options in developing nations. But the reality could be more complicated. A team at the Joint Global Change Research Institute worked with Victor and others to investigate the risks of making investments in developing countries due to political instability and the relatively poor quality of many public institutions there. Their model showed[4] that investors would probably shun developing countries and pour money into developed ones, driving up costs and making it harder to curb rapidly rising emissions in developing nations.

"The models have taught us that with unrealistic assumptions anything is possible, and with realistic assumptions it will be very hard to cut emissions to meet goals like 2 degrees," Victor says. "That's an important result because it forces — or should force — some sobriety about what can be achieved."

One message that modellers have delivered quite clearly is that without collective and aggressive action by all countries, costs invariably increase, and the chance of hitting the 2 °C goal plummets. This is precisely the situation heading into the Paris summit. Most countries, and all of the major greenhouse-gas emitters, have submitted pledges to reduce their emissions, but these vary widely in ambition.

As it stands, the world is on a path to nearly 3 °C of warming by the end of the century, and even that assumes substantial emissions reductions in the future. If nations do not go beyond their Paris pledges, the world could be on track to use up its 2 °C carbon budget as early as 2032. If the models are correct, world leaders may have to either accept extra warming or plan for a Herculean negative-emissions campaign. In the event that they choose the latter — and succeed — the entire debate will change.

"It's a completely different game," says Nebojsa Nakicenovic, an economic modeller and deputy director-general of the International Institute for Applied Systems Analysis in Laxenburg, Austria. "If that is technically possible, then we could also go below 2 degrees."

Fast-forward to 2100 once more. The bioenergy industry is now one of the largest and most powerful on Earth. People are pulling roughly as much CO_2 out of the atmosphere as they were emitting at the time of the historic Paris conference. Humanity has asserted control over the atmosphere, and governments face a new and difficult question at the 108th anniversary of the UN climate convention: how low should they set the global thermostat? ■

Jeff Tollefson *writes for* Nature *from New York.*

1. Allen, M. R. *et al. Nature* **458,** 1163–1166 (2009).
2. Kato, E. & Yamagata, Y. *Earth's Future* **2,** 421–439 (2014).
3. Fuss, S. *et al. Nature Clim. Change* **4,** 850–853 (2014).
4. Iyer, G. C. *et al. Nature Clim. Change* **5,** 436–440 (2015).

COMMENT

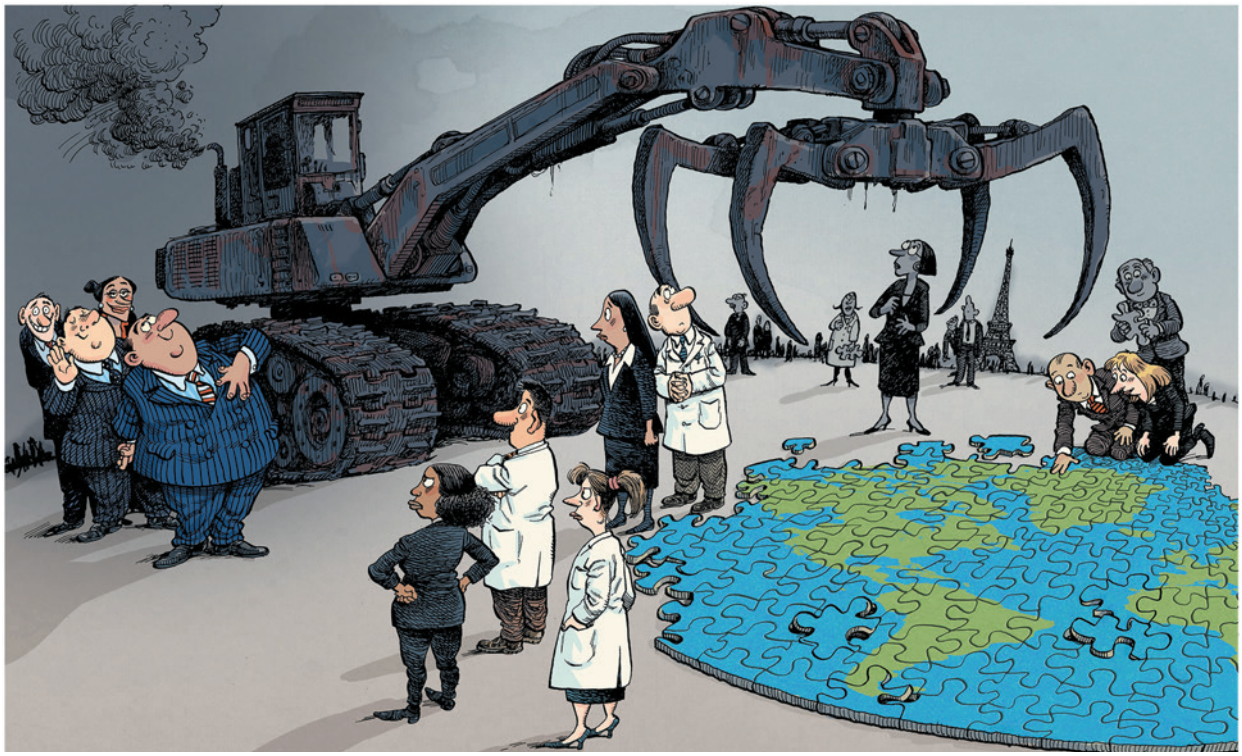


ILLUSTRATION BY DAVID PARKINS

After the talks

The real business of decarbonization begins after an agreement is signed at the Paris climate conference, argue **David G. Victor** and **James P. Leape**.

After years of failure to craft global agreements on climate change, the upcoming United Nations Paris Climate Conference is likely to turn a corner. Diplomats have drafted a workable text that will probably be adopted. Businesses and environmental groups are engaged in the process in unprecedented ways.

Governments, development banks and foundations are raising funds to help the poorest countries to pay for cutting emissions and prepare for a changing climate[1] — the main sticking point in 2009, when the last big climate conference, in Copenhagen, ended in disarray. The UN and the French hosts have a sophisticated agenda to bring all these efforts together. Even religious leaders have spoken mightily of the dangers of unchecked climate change.

Good news from the Paris meetings will build confidence, a crucial ingredient for effective international cooperation. Governments and firms will invest in a future with lower emissions if they think that others will do the same[2]. Agreement will demonstrate the viability of a new, flexible 'bottom-up' mode for climate diplomacy — based on national pledges that accommodate different preferences and capabilities. By contrast, the rigid targets and timetables of the Kyoto Protocol appealed to few of the world's emitters.

Yet a dose of sobriety is also needed. Agreements are feasible now only because diplomats are postponing the thorniest problems, such as how to hold nations accountable. Business engagement may prove ephemeral when the spotlight shifts. Good news about climate finance is possible now because the blend of public funding (which is hard to mobilize and spend effectively) and private money (which is abundant but often rarely focused on global goals) is vague.

Whether the Paris conference will succeed depends on what unfolds afterwards. Diplomats will have much to do until 2020, when the main accords take full effect. Civil society — notably business — must shift from making bold promises to cutting emissions. Governments and business must build ▶

and invest in review and accountability mechanisms to ensure that they are keeping their promises — an area in which non-governmental organizations (NGOs) have a crucial role. And scientists must pursue research that is directly relevant to policy-making, as well as assessing the underlying causes and impacts of climate change.

ENGAGE BUSINESS

Keeping business on board will be the most important challenge. It is easy for companies to make commitments when the world's media and political leaders are watching. It is harder to implement changes when cut-throat competition makes it risky to invest in more expensive but less polluting technologies and practices.

The most striking example of business engagement is the pledges that many firms and governments are making to cut deforestation[3]. In 2010, the Consumer Goods Forum (comprising the largest retailers and consumer-products companies) announced that its members would eliminate deforestation from their supply chains, notably for palm oil, soya, beef, timber and pulp. More than 300 companies have followed suit (see www.supply-change.org). Leading producers and traders of palm oil in Indonesia — which accounts for half of the world's supply — have promised to stop converting forest or peat lands[4]. Palm oil is a main culprit in the fires that have spread a choking haze across the region since August, afflicting more than 40 million people and often causing daily emissions of greenhouse gases that surpass those of the United States.

It is far from assured that these pledges will result in lasting changes in the complex supply chains — from how the land is managed, to the produced oil and finally to consumer products. There are already signs of trouble. Most businesses pledge to become more sustainable following pressure from NGOs[5]. (One of us, J.P.L., led WWF International for nine years, during which time the organization was centrally involved in many such efforts.) Firms fear consumer backlash if their products are tied to environmental destruction[6] (see go.nature.com/5l8yjm). After the Paris meetings, chief executives will need to activate changes through the ranks of their organizations and suppliers; NGOs will need both to maintain the pressure for action and to work with companies to secure broader reforms in major producing countries.

Shifting whole industries into more sustainable modes of production requires collaboration between government, business and civil society. Economic incentives must be rewired so that no firm can gain an advantage by, for example, continuing to destroy forest. Solutions will vary by country and locality, but common threads include better governance — laws, fiscal regimes, property rights and public administration — and investment in helping countries, communities and small producers to make the transition to sustainability.

Brazil has shown what is possible. Between 1995 and 2005, forest loss in the Brazilian Amazon averaged 19,500 square kilometres per year — roughly the area of Israel[7]. By 2013, that rate had been cut by 70%, even as beef and soya production continued to grow. A combination of measures was applied: corporate commitments coupled with strong laws, satellite surveillance and robust enforcement, restrictions on access to credit for farms and ranches in counties with high deforestation, the creation of protected areas and indigenous reserves, and improvements in land tenure and governance. Brazil's federal government worked closely with the beef and soya industries, NGOs and international partners. In 2008, for example, Norway committed US$1 billion to Brazil because it wanted to demonstrate practical new ways to protect forests globally. Even so, Brazil's progress is fragile — deforestation in the Amazon has increased over the past 18 months.

"It is easy for companies to make commitments when the world's media and political leaders are watching."

Beyond forestry, industry's commitment to reducing emissions is mixed. The three dozen firms and governments that account for 40% of the methane released from oil and gas production have pledged to eliminate those emissions by 2030, for example (see go.nature.com/beuw2z). Details on how this pledge will be monitored are scarce, as is a plan to extend the pledge to the rest of the global industry.

Business is, mostly, still waiting to see whether the Paris conference will turn out to be a watershed. Governments are looking for signs that industry can cut emissions at acceptable cost and are sceptical that competing nations will take action. For all the good will in Paris, this chicken-or-egg problem looms large — it explains why climate policy requires international cooperation, and why so little progress has been made over the past 25 years. Governments must grapple with huge unknowns about what mitigation will cost and whether other countries will honour their commitments[8]. Until confidence in international cooperation grows, politicians and business leaders will talk big but deliver small[2].

NEW DIPLOMACY

Optimism about Paris is partly rooted in a new bottom-up bargaining system whose flexibility, in theory, is suited to crafting policy in areas in which cooperation is essential but countries are unsure about what is feasible[2,8]. National pledges — in diplomatic jargon, 'intended nationally determined contributions' (INDCs) — allow governments to align their commitments with national priorities. This approach has elicited firm commitments — notably from countries such as the United States, China and India, which are skittish about inflexible international legal commitments yet willing to do their part for the global whole. China's pledges, for example, will help to slow global warming while serving the country's pressing concerns about reducing air pollution and achieving energy security.

Pledge systems also bring dangers. The current round of INDCs is thin on content; some countries have failed to supply any reports, and industry has been largely absent from the process. Unless the pledging system is improved, it could become a licence to do nothing. This is why earlier schemes have yielded little practical action — as with the Asia-Pacific Partnership on Clean Development and Climate created in 2005 by then-US President George W. Bush after the United States refused to ratify the Kyoto Protocol. Pledges must offer enough detail and transparency for diplomats to link national efforts into more-ambitious, collective agreements in the future. A priority after Paris will be to develop stricter standards for national pledges as well as robust systems for review.

ROAD AHEAD

Only so much can be achieved within the UN system — in which consensus is usually required and it is easy for reluctant nations to block progress. Countries and firms will need to find ways to work in smaller, focused and more practical groups — in tandem with the broader global objectives[8]. Doing this cannot rest on altruism — it requires attention to self-interest and, as the palm-oil example shows, putting pressure on governments and firms to rethink their self-interest.

Countries that want this new flexible system to work should volunteer to do more — for example, by offering their INDCs for reform and review. The United States and China should offer their own bilateral climate accord, made in November last year — which pledged emissions curbs and efforts to conduct joint research on new technologies — to independent scrutiny, such as by the Organisation for Economic Cooperation and Development or the World Bank. With a huge stake in showing the effectiveness of the pledging process, these two countries must bear the burden of proof[8].

Firms, too, must recognize that their efforts will be believed only with transparency and public accountability. Failure to demonstrate that corporate pledges are

leading to tangible action will lead to demands after Paris for more onerous and costly regulation. Industry pledges should be reviewed alongside government commitments — and leading firms that have the most to gain from this new system of governance should invest in the needed independent reviews. NGOs have a key role in holding companies to account, assessing to what degree stated reductions are real (with no double counting) and identifying where extra effort is needed.

For academics, this world of bottom-up diplomacy demands new skills. Periodic global assessments of the state of the science and gaps between what governments and firms pledge and what the planet needs for protection will still be needed. Equally urgent is interdisciplinary research predicting how these messy, decentralized systems of governance will function. Scientists, including social scientists, will need to look, together, at how societies develop and implement policy reforms while assessing what works so that research is more informative for policymakers.

Sceptics will see that messy reality on display at the Paris conference and declare that the event has failed to deliver on widely discussed goals such as stopping warming at 2 °C above preindustrial levels. The better metric is whether Paris engages a growing share of industry and governments in the climate task. When the meetings in Paris are done, the real business of decarbonization must begin. ■

David G. Victor *is professor of international relations at the School of Global Policy and Strategy, University of California, San Diego, California, USA.* **James P. Leape** *is consulting professor in the School of Earth, Energy and Environmental Sciences & Woods Institute for the Environment, Stanford University, California, USA.* **D.G.V.** *and* **J.P.L.** *are also on the World Economic Forum's Global Agenda Council on Governance for Sustainability.*
e-mail: david.victor@ucsd.edu; jleape@stanford.edu

1. OECD. *Climate Finance in 2013–14 and the USD 100 Billion Goal* (OECD, 2015).
2. Victor, D. G. *Global Warming Gridlock* (Cambridge Univ. Press, 2011).
3. Hsu, A. *et al. Nature Clim. Change* **5,** 501–503 (2015).
4. Carlson, K. M. & Curran, L. M. *Carbon Mgmt* **4,** 347–349 (2014).
5. Potoski, M. & Prakash A. (eds) *Voluntary Programs: A Club Theory Perspective* (MIT Press, 2009).
6. Overdevest, C. & Zeitlin, J. *Regul. Gov.* **8,** 22–48 (2014).
7. Nepstad, D. *et al. Science* **344,** 1118–1123 (2014).
8. Sabel, C. F. & Victor, D. G. *Clim. Change* http://dx.doi.org/10.1007/s10584-015-1507-y (2015).

PHIL WILLS/ALAMY

Corrosion costs around US$4 trillion a year globally.

Share corrosion data

To prevent disasters, **Xiaogang Li** and colleagues call for open data infrastructures to collate information on materials failures.

In November 2013, an oil pipeline in the Chinese city of Qingdao exploded, killing 62 people and wounding 136. Eight months later, a similar explosion in Kaohsiung caused 32 deaths and 321 injuries. The pipelines were made of steel of the same specification and they failed after two decades of use in similar environments. The cause was corrosion — the degradation of a material by a chemical or electrochemical reaction with its environment.

Such disasters are common: each square kilometre of any Chinese city hosts more than 30 kilometres of buried pipes, creating tangled networks of oil and gas lines, water mains and electrical and telecommunications cables. Corrosion is costly, too. According to a US survey, corrosion costs six cents for every dollar of gross domestic product in the United States[1]. Globally, that amounts to more than US$4 trillion a year — equivalent to damages from 40 Hurricane Katrinas. Half of that cost is in corrosion prevention and control, the other half in damages and lost productivity.

A lack of knowledge hinders our ability to prevent failures. Degradation of underground pipes, for example, is influenced by the compositions, microstructures and designs of materials, as well as by a raft of environmental conditions such as soil oxygen level, humidity, salinity, pH, temperature and biological organisms.

Many industries, including oil, gas, marine and nuclear, collect corrosion data to identify risks, predict the service lives of components and control corrosion. Most of these data are proprietary, and best practices are rarely shared. Oil spills, bridge collapses and other disasters continue to occur.

Demand for knowledge about corrosion is growing, with the increasing use of advanced materials in medical devices, biosensors, fuel cells, batteries, solar panels and microelectronics. Corrosion is the main restriction on many nanotechnology applications.

Efforts to make materials data accessible, such as the Materials Genome Initiative (MGI), focus on 'births' rather than ▶

'deaths' of materials. Online platforms for sharing corrosion data are badly needed. Access to a large volume and variety of corrosion information that researchers could probe with data mining and modelling tools would improve forecasts of corrosion failures and anticorrosion designs.

COMPLEX PROCESSES

The biggest challenge in corrosion research is predicting accurately how materials will degrade in a given environment[2]. It requires full knowledge of all relevant factors and their interactions. Yet precise models for mechanisms are lacking. Forecasting problems is impossible without historical data about materials failures under various conditions. And field performances cannot be judged in laboratories when environmental parameters are unknown.

Corrosion data are hard to collect. Damage may take years or decades to accumulate and any project tracks only a handful of contributing factors. Data sets need to be combined. For example, early studies of marine corrosion (occuring, for instance, on oil-drilling platforms) were unreliable because they considered only physiochemical processes (those involving pH, dissolved oxygen and temperature) and not the effects of organisms living in seawater. The inclusion of genomic data has now improved the models.

Corrosion depends on local conditions. Steel structures that last for decades in dry parts of inland China fail within months in humid and salty coastal areas of southeast Asia. Protective polymer coatings that work for years at northern latitudes can degrade in weeks near the Equator, where heat and greater doses of ultraviolet radiation break chemical bonds more quickly. Inferring general corrosion knowledge — such as how particular steels are affected by humidity, salt or air pollution — requires combining studies from many diverse environments. One worldwide survey of weathering steel, for example, reviewed exposure test results for up to 22 years from 108 sites in 22 countries[3].

With global trade increasing, the oil and gas, construction, car, electronics and other industries have called for corrosion data to be shared between countries to ensure the quality and safety of their products. Millions of cars worldwide have been recalled in the past few years owing to unforeseen corrosion problems arising in destination countries. China's 2013 'Belt and Road' initiative, which promotes industrial ties with countries along the Silk Road economic belt between China and the West, raises unprecedented challenges. Rapid corrosion assessment, materials selection and design will be needed as billion-dollar construction, transport, energy and telecommunications projects begin in Asia, Africa and Europe.

Advanced materials present entirely new corrosion problems. For example, the electrochemical stabilities of noble metals such as platinum and gold fall sharply as their dimensions decrease to nanometre scales. Corrosion of platinum nanoparticles remains a roadblock limiting the lifetime of platinum-based catalysts for fuel cells.

Corrosion scientists have been slower than their materials-science peers to recognize the need for data sharing. Several large materials-data repositories built by US governmental agencies under the auspices of the MGI house basic physical, chemical and microstructure data for materials, but not corrosion data. Yet none of the advanced materials promised by the MGI will be practical without considering their environmental stability and durability.

> *"Forecasting problems is impossible without historical data about material failures."*

DATA REPOSITORIES

Open data infrastructures should be set up to house corrosion data in various countries, industries and applications. By using the same standardized formats for data and metadata, the data can be connected and eventually amount to a global system, possibly linked to the MGI.

Governments should take the lead. For example, the Chinese government has invested nearly 200 million yuan (US$30 million) since 2006 on a platform for sharing corrosion data from 30 field-testing stations covering standard materials in environments (air, soil and water) typical of different parts of the country. Other nations, industries and interest groups should establish similar data infrastructures for corrosion in other regions and sectors.

Efforts need to be coordinated to collect corrosion data that are relevant to urgent or emerging challenges, such as alternative energy and nanotechnology. For instance, the US Department of Energy has partnered with the MGI to build materials data repositories to help to speed up the development of alternative clean-energy sources.

Funding agencies should incentivize the sharing of corrosion data about advanced materials and emerging technologies, for example, by demanding it in research grants and supporting the costs of publishing in open-access journals. Corrosion-science societies should learn from general materials-science societies (such as Materials Research Society, the Minerals, Metals & Materials Society and ASM International) and convene experts to establish data-sharing best practices and guidelines.

Industry involvement can be encouraged through partnerships with academia. Companies would save research and development costs in return for contributing data to repositories. Because corrosion concerns maintenance and safety rather than industrial competition, businesses should be willing to share such data. Data consortia can be formed to identify common topics of priority and jointly develop benchmark solutions, just as industrial standards are agreed.

More-powerful tools need to be developed for data capturing, management, mining, modelling and simulation — the integration of which we term corrosion big data and informatics[4]. Advanced monitoring technologies require 'big data' analytics. For instance, robots (known as 'smart pigs') carrying hundreds of sensors deployed to inspect the walls of pipelines can collect 1 terabyte of data in one run. Highly accurate corrosion simulations could partially or completely replace the time-consuming, environmentally unfriendly, complicated and expensive experimental corrosion tests. For example, quantum chemical simulations are heavily used to evaluate the molecular structures and electronic properties of corrosion inhibitors[5].

If corrosion data is shared, everyone will benefit from the greater understanding that results. ■

Xiaogang Li *is professor at the Key Laboratory for Corrosion and Protection of the Ministry of Education, Institute of Advanced Materials & Technology, University of Science and Technology Beijing, Beijing, China, and is at the Ningbo Institute of Material Technology & Engineering, Chinese Academy of Sciences, Ningbo, Zhejiang, China.* **Dawei Zhang**, **Zhiyong Liu**, **Zhong Li**, **Cuiwei Du** *and* **Chaofang Dong** *are at the Key Laboratory for Corrosion and Protection of the Ministry of Education, Institute of Advanced Materials & Technology, University of Science and Technology Beijing, Beijing, China.*
e-mail: dzhang@ustb.edu.cn

1. Koch, G., Brongers, M., Thompson, N., Virmani, Y. & Payer, J. *Corrosion Cost and Preventive Strategies in the United States* (NACE International, 2002).
2. Duquette, D. *et al. Research Opportunities in Corrosion Science and Engineering* (National Academies Press, 2011).
3. Morcillo, M., Chico, B., Díaz, I., Cano, H. & de la Fuente, D. *Corrosion Science* **77,** 6–24 (2013).
4. Li, X. *Informatics for Materials Corrosion and Protection: The Fundamentals and Applications of the Materials Genome Initiative in Corrosion and Protection* (Chinese Chemical Industry Press, 2014) (in Chinese).
5. Taylor, C., Chandra, A., Vera, J. & Sridhar, N. *Faraday Discussions* **180,** 459–477 (2015).

CORRECTION

The Comment article 'Einstein was no lone genius' (M. Janssen and J. Renn *Nature* **527,** 298–300; 2015) wrongly stated the dates during which Albert Einstein studied at the Swiss Federal Polytechnical School in Zurich. He was there between 1896 and 1900.

NASA

The first iconic image of Earth from space sparked awareness of planetary boundaries.

SUSTAINABILITY

The launch of Spaceship Earth

Adam Rome revisits five prescient classics that first made sustainability a public issue in the 1960s and 1970s.

In 1969, in a book-length essay entitled *Operating Manual for Spaceship Earth*, the inventor and polymath Buckminster Fuller offered a striking metaphor for a new ideal of planetary management. Although Earth did not come with instructions, our spaceship had built-in safety features that had kept us going. Still, our pilot errors were catching up with us: we had been so "misusing, abusing, and polluting" the planet, Fuller argued, that it might need to be renamed "Poluto". That way lay humanity's oblivion. But if we discovered how our spaceship worked — if we learned to make the best use of our incredible ingenuity — we might become "comprehensively and sustainably successful".

Like everything Fuller wrote, *Operating Manual for Spaceship Earth* was idiosyncratic, at once arresting and fanciful. But many of the book's basic ideas were in the air at the time. Between roughly 1965 and 1975, the challenge of sustaining civilization inspired a shelf-full of influential books. They had a freshness, urgency and breadth that are hard to credit today, and they are still remarkably relevant. Now that sustainability as a concept has become dulled by overuse, they return our eyes to the prize.

These seminal studies built on earlier fears. Fairfield Osborn's *Our Plundered Planet* (Little, Brown) and William Vogt's *Road to Survival* (W. Sloane Associates), both published in 1948, warned that uncontrolled population growth and resource depletion would lead to calamity. But the situation seemed even more precarious by 1970, when the first Earth Day was celebrated across the United States. The human impact on the planet had exploded after the Second World War, and scientific advances had led to greater understanding of the threat from those impacts. For the first time, many realized, we had the potential to disrupt or even destroy the planet's life-support systems. The sense of environmental crisis was exacerbated by the social and political turmoil of the period.

What would be required for humanity to continue to thrive? To tackle so huge a question required intellectual audacity, and the authors of the pioneering books on sustainability were all big-picture, interdisciplinary thinkers par excellence. Economist Kenneth Boulding — author of *The Meaning of the Twentieth Century* (1964) — thought historically and philosophically. Biologist Barry Commoner felt compelled to study political economy, as his 1971 *The Closing Circle* shows. Fuller considered himself a futurist. The authors of the 1972 *The Limits to Growth* — Donella Meadows, Dennis Meadows, Jørgen Randers and William Behrens — meshed environmental science with systems analysis. Barbara Ward was a journalist, economist and adviser to world leaders who collaborated with Pulitzer-prize-winning microbiologist René Dubos on *Only One Earth* (1972).

The Meaning of the Twentieth Century is no longer well known, yet Boulding was key in framing the issue of sustainability. He made clear that the world that he hoped to sustain did not yet exist: humanity was in the middle of a "great transition" from an agricultural species to a thoroughly industrial one. In Boulding's view, this transition was fraught with peril and sure to be wrenching. It might be derailed by nuclear war or uncontrolled population growth, and might fail if we misused natural resources, especially fossil fuels. To succeed, we needed to create "a stable, closed-cycle, high-level technology" that would not pollute or require exhaustible materials. (He expanded on that in an often-reprinted 1966 essay, 'The economics of the coming spaceship Earth'.) But developing new technology was not the heart of Boulding's prescription. He argued that a sustainable future would require countless "social inventions", from new aesthetics to better methods of resolving disputes. "The unfinished tasks of the great transition are so enormous," he concluded, "that there is hardly anyone who cannot find a role to ▶

The Meaning of the Twentieth Century: The Great Transition
KENNETH E. BOULDING
Harper and Row: 1964.

Operating Manual For Spaceship Earth
R. BUCKMINSTER FULLER
Southern Illinois University Press: 1969.

The Closing Circle: Nature, Man, and Technology
BARRY COMMONER
Knopf: 1971.

The Limits to Growth: A Report for the Club of Rome's Project on the Predicament of Mankind
DONELLA H. MEADOWS, DENNIS L. MEADOWS, JØRGEN RANDERS, AND WILLIAM W. BEHRENS III
Universe: 1972.

Only One Earth: The Care and Maintenance of a Small Planet
BARBARA WARD AND RENÉ DUBOS
W. W. Norton: 1972.

HANS NAMUTH/SPL

YUTAKA NAGATA/UN PHOTO

Inventor Buckminster Fuller (top) approached sustainability as a design challenge; economist Barbara Ward (bottom) prompted the United Nations to integrate social and environmental issues.

play in the process." That is still ever true now: dealing with climate change requires a host of skills.

Fuller broke new ground by defining sustainability as a design challenge. Already famous for inventions such as the strong, lightweight, geodesic dome, he wrote exuberantly about the need for an "industrial retooling revolution": to achieve lasting affluence, we must learn to do more with less. Like Boulding, Fuller argued that we needed to treat fossil fuels as a short-term expedient while we worked out how to fashion a sustainable future. For a reader today, the insights of Fuller's work are not enough to make up for the idiosyncrasies of his language and argument. William McDonough and Michael Braungart's *Cradle to Cradle* (North Point, 2002) would be a much better introduction to sustainable design. But in 1969, Fuller's work seemed thrilling, and his *Operating Manual* became a bible for people keen to invent eco-efficient ways of providing energy, building things and managing wastes.

Commoner's *The Closing Circle* laid the foundation for industrial ecology. Particularly in the postwar decades, Commoner argued, the industrialized world had come to rely on a host of "ecologically faulty" technologies, from nuclear power to chemical pesticides. The technologies of the future needed instead to accord with four basic principles, which he defined as laws of ecology: "Everything is connected to everything else", "Everything must go somewhere", "Nature knows best" and "There is no such thing as a free lunch".

For Commoner, however, the ultimate problem was economic and political, not technological. Discussing the economic meaning of ecology, he argued that the private-enterprise system had serious flaws. Businesses had powerful incentives to produce new products that did more environmental harm than the products they replaced. They did not need to account for "biological capital", and they did not pay the full costs of production, which included pollution. In the decades since *The Closing Circle* appeared, making capitalism greener has become a major concern of economists, business-school professors, entrepreneurs, corporate executives and activists, yet much of Commoner's critique still holds.

The Limits to Growth asked — heretically — whether humans could continue indefinitely to make ever greater demands on Earth. The authors used computer modelling to explore the interactions between population growth, resource demand, industrialization, food production and pollution. They did not forecast the future, although commentators ever since have debated whether their 'predictions' were right; instead, they extrapolated. If present trends continued, the authors wrote, humanity would hit the wall "sometime within the next hundred years". They hoped that people would avert a breakdown, but stated repeatedly that they could not model the social, political and cultural factors that might alter trends. They did consider whether technology could be a magic bullet, and the results were shocking. Even when they allowed for the technological progress that greatly increased the availability of resources and reduced the amount of pollution, the result was still collapse — just farther down the road. Innovation alone could not lead to a sustainable economy. We needed

a fundamental shift in values.

The Limits to Growth was an international sensation, selling over 12 million copies in more than 30 languages. Meadows, Meadows and Randers updated the analysis in 1993 and again in 2004, and the question of limits still prompts vigorous debate. Johan Rockström and Mattias Klum's *Big World, Small Planet* (Yale University Press, 2015) and Donald Worster's *Shrinking the Earth* (Oxford University Press, 2016) are just two of the many books now probing the problem of growth.

Ward and Dubos's *Only One Earth*, written to accompany the 1972 United Nations Conference on the Human Environment, added an international perspective to the sustainability discussion. Ward had travelled the globe as an expert on economic development. A preliminary draft of the book was circulated for comment to scientific, business and intellectual leaders from 58 countries, and the result is worth reading just for the summary of their responses, which made clear that people around the world held very different views about environmental issues. A European respondent argued for a retreat from industrialization, for example, whereas an Asian statesman wrote that developing nations could not afford "dreams of landscapes innocent of chimney stacks".

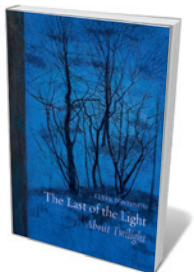
For Ward and Dubos, any effort to ensure the survival of humanity had to bridge the tremendous gap between developed and developing nations. Although they didn't use the phrase 'sustainable development', they offered a path-breaking analysis of the challenge of raising living standards for the poor without degrading the environment. At the same time, they called for the affluent to take off their blinkers. Well-to-do nations needed to acknowledge the damage that they were doing to the biosphere — and to accept that their fate was inseparable from the prospects of the rest of the world. Because many environmental threats were global, Ward and Dubos concluded, "planetary interdependence" had to become a moral and political reality, not just "a hard and inescapable scientific fact". The UN Paris Climate Change Conference starting this month will be a test of how close we are to meeting that aim.

Read together, the books of this charged decade demonstrate that building a sustainable civilization is multidimensional. It sweeps everything in: science and technology, politics, economics, social relationships, ethics. We cannot advance in a straight line. We need to approach the goal from many directions, with flexibility and tenacity. ■

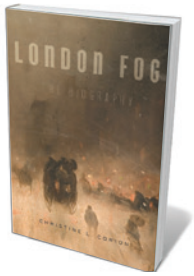
Adam Rome *is a professor of history and English and the Unidel Helen Gouldner Chair for the Environment at the University of Delaware in Newark. His latest book is* The Genius of Earth Day.
e-mail: arome@udel.edu

Books in brief

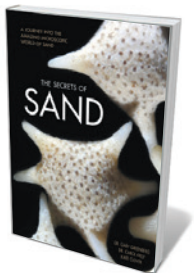
The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter
Joseph Henrich PRINCETON UNIVERSITY PRESS (2015)
The force propelling *Homo sapiens* down its unique evolutionary pathway is "culture-gene coevolution", avers anthropologist (and aerospace engineer) Joseph Henrich. Over time, he posits, the need to acquire "adaptive cultural information" expanded the human brain, and societies' "collective brains" in turn shaped human culture. Integrating insights from cognitive psychology, experimental economics, history and ethnography, this limber and lucid study concludes that we face a major transition into a new type of animal.



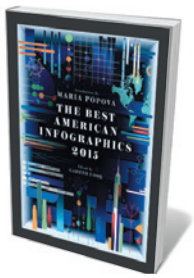
The Last of the Light: About Twilight
Peter Davidson REAKTION (2015)
Cultural historian Peter Davidson enters the twilight zone, tracing the crepuscular in science, psychology, history and the arts. Considering the 60th parallel north, around which "long evenings and protracted sunsets stretch", Davidson probes aspects of this transitional state, including visual perception during the stages of twilight (civil, nautical and astronomical); dusk as a metaphor for crisis in Charles Dickens's *Bleak House*; the proliferation of gilt and mirrors in the murky pre-electric era; and the poet Gerard Manley Hopkins' observations of anti-crepuscular rays, published in *Nature*.



London Fog: The Biography
Christine L. Corton BELKNAP (2015)
London's 'pea-soupers' — opaque, yellowish smogs — were an environmental catastrophe, a cloak for nefarious activities and an artistic inspiration. An odiferous wig of soot from coal fires, sulfur dioxide and mist settled regularly over the city from the 1840s to the 1960s. In this richly nuanced history, scholar Christine Corton takes us from polymath Robert Hooke spotting a pall of smoke over London in 1676 through the killer fogs that felled zoo animals, spurred crime and caused traffic accidents, and that ultimately galvanized scientists and the government to craft the 1956 Clean Air Act.



The Secrets of Sand: A Journey into the Amazing Microscopic World of Sand
Gary Greenberg, Carol Kiely and Kate Clover VOYAGEUR (2015)
Beachcombers take heed: the real treasure is stuck to your soles. Sand — as cell biologist Gary Greenberg, microscopist Carol Kiely and science curator Kate Clover show in this delightful coffee-table book — is dazzling, from star-shaped forams to egg-like ooids. To photograph these minuscule jewels rock-polished by wind and surf, Greenberg used 3D microscopes and smart lighting. A stunning extra are images of the lunar dust particles that Kiely studies, including glassy spherules from extinct fire-fountain volcanoes.



The Best American Infographics 2015
Gareth Cook and Maria Popova MARINER (2015)
Another year, another superb volume in this infographics series edited by journalist Gareth Cook; cultural curator Maria Popova (of blog 'Brain Pickings') guest-introduces. 'What Do Americans Speak?' (*Slate*, 13 May 2014) offers an eye-popping map showing the third most commonly spoken language in each US state — in Michigan, that is Arabic — and *Nature*'s own 'Born Here, Died There' (*Nature* http://doi.org/8xg; 2014) explores dynamic patterns in cultural history through an elegant animation. **Barbara Kiser**

Correspondence

Gene editing: heed disability views

CRISPR–Cas9 is a gene-editing tool of great potential, although not necessarily from a disability-rights perspective (see D. J. H. Mathews *et al. Nature* **527,** 159–161; 2015). People with disabilities are, in my view, unlikely to be queuing up for genetic modification: their priority is to combat discrimination and prejudice.

To 'fix' a genetic variation that causes a rare disease may seem an obvious act of beneficence. But such intervention assumes that there is robust consensus about the boundaries between normal variation and disability. Contrary to the prevailing assumption, most people with disabilities report a quality of life that is equivalent to that of non-disabled people (G. L. Albrecht and P. J. Devlieger *Soc. Sci. Med.* **48,** 977–988; 1999).

The UK Nuffield Council on Bioethics is deliberating the ethical and social dimensions of CRISPR. International guidelines are urgently needed (*Nature* **526,** 310–311; 2015), and the voices of people living with illness and impairment need to be heard.

Tom Shakespeare *University of East Anglia, Norwich, UK.*
tom.shakespeare@uea.ac.uk

Gene editing: govern ability expectations

From a disability-rights viewpoint, problems that have dogged the debate on human genetic modification (see go.nature.com/6wb45k) also pervade your curtain-raiser to the US National Academies of Sciences, Engineering and Medicine conference (see D. J. H. Mathews *et al. Nature* **527,** 159–161; 2015). The authors' portrayal of the public as a passive recipient of 'wisdom' from 'experts' goes against healthy discourse on responsible research and governance.

The disability-rights community has a history of disagreement with such experts (including authorities, scientists and clinicians) over their perception of people with disabilities. This is summarized as 'ableism', a view that disability is an abnormality instead of a feature of human diversity. It can lead to flawed 'solutions' and disempower those affected (see G. Wolbring *J. Crit. Anim. Stud.* **12,** 118–141; 2014).

"It is time to collectively make decisions about the kind of world we want to live in," write Mathews and colleagues. This discussion should include ability expectations and how they should be governed.

Gregor Wolbring *University of Calgary, Alberta, Canada.*
gwolbrin@ucalgary.ca

Gene editing: survey invites opinions

As the US National Academies of Sciences, Engineering and Medicine summit on the regulation of CRISPR–Cas9 gene-editing tools gets under way, we invite readers to contribute their opinions about this technology and its use to a survey at go.nature.com/eyowaf.

Public engagement in decisions about applications of science and technology that affect society is essential. The summit is, to a degree, modelled on the 1975 Asilomar Conference on the potential biohazards of recombinant DNA (see *Nature* http://doi.org/899; 2015). It must not make the same mistake of being held behind closed doors.

As one survey contributor remarks, it may be impossible "to get this [CRISPR–Cas9] genie back into the bottle". So when it comes to wishes for the genie, those of both scientists and the public must be considered.

Silvia Camporesi, Lara Marks *King's College London, UK.*
silvia.1.camporesi@kcl.ac.uk

Climate change also creates expatriates

I visited the island of Tuvalu in the Pacific Ocean three decades ago as the environmental assessor for an aid-funded engineering consultancy. Pollution of the freshwater lens and scavenging of protective shoreline coral rubble for construction were problems even then. As you note (see *Nature* **526,** 624–627; 2015), these may drive exodus sooner than rising sea levels.

Nobody likes to be forced out of their home. But small oceanic nations hold a valuable asset: sovereignty. Tuvalu already profits from its own Internet domain (.tv), and sovereign nations have United Nations votes, which are effectively on the market. They can operate attractive tax regimes. They can declare marine reserves and sell rights to fisheries, seabed mining or reef tourism. All of these make money, and it does not have to be divided between many people. They can all be done even if nobody lives there in person. Citizens of such small island nations could thus become well-off expatriates, as well as refugees.

Ralf Buckley *Griffith University, Gold Coast, Australia.*
r.buckley@griffith.edu.au

Crowdfunding not fit for clinical trials

Crowdfunding can raise money quickly and with minimal bureaucracy. But it should not be considered as a way to finance clinical trials because of potential ethical implications.

One problem is that funding recipients are not accountable to the public because crowdfunding is unregulated. Another is that there is no setting of research priorities, so crowdfunded clinical trials may not be the most important or widely applicable ones. And media tactics could attract emotional donations, for example by generating false expectations of a 'cure'. Moreover, an inconclusive or negative outcome could erode public trust.

By contrast, the mainstream funding process for clinical trials takes into account disease prevalence, morbidity and mortality, justice and utility. Crowdfunding for clinical trials should be similarly regulated to mitigate its potential risks.

Phaik Yeong Cheah *University of Oxford, UK.*
phaikyeong@tropmedres.ac

Lessons from EPA on tracking pollutants

In our opinion, China could learn from the success of the US Environmental Protection Agency (EPA) in providing open-access environmental information to the public. This would enhance the credibility of government decisions.

The EPA's Toxics Release Inventory programme, in partnership with state agencies, collects data from enterprises that must report emissions. It subjects this information to quality-assurance reviews, trends analysis and error correction, as well as making it publicly available. This evaluation of the entire information-flow process increases transparency and accountability.

Using a comparable holistic approach, China's Ministry of Environmental Protection could develop a secure access point for ministries and agencies and a web portal for public access. A designated group might set quality standards and policies for handling such information, akin to the EPA's Office of Environmental Information.

Bo Zhang *Information Center, Ministry of Environmental Protection, Beijing, China.*
Wayne S. Davis *EPA, Washington DC, USA.*
zhangbo@mep.gov.cn

GENOMICS

Acorn worms in a nutshell

The genome sequences of two members of the hemichordate group of marine invertebrates bring the evolution of their relatives, including vertebrates, into sharper focus. SEE ARTICLE P.459

CASEY W. DUNN

By examining the similarities and differences among the genomes of living organisms, we can reconstruct features of the genomes of long-dead ancestors. Such reconstructions provide insight into patterns of genome diversity and how organisms evolved through the gain, loss and modification of genomic features. The greater the number of sequenced genomes from living organisms, and the broader their distribution across the tree of life, the better is our view of these ancestral genomes. However, although hundreds of animal genomes have been published in recent decades, the vast majority are from only two groups: vertebrates and arthropods. Simakov and colleagues' publication[1] in this issue (page 459) of genome sequences for two species from a group of invertebrates known as hemichordates takes the sampling of animal genomes an important step forward.

Hemichordates are exclusively marine animals. The adults live on the ocean bottom, whereas the larvae are free-swimming. There are about 130 described species[2], which are divided into 2 groups. The pterobranchs, of which there are around 20 species, are small animals (up to about 5 millimetres long) that form colonies of asexually produced clones attached to a central disk by fleshy tethers[3]. The animals live in a tube network that they secrete. Just as birds were found to be 'living dinosaurs' — a group that had been thought extinct — pterobranchs are living graptolites, animals that are abundant in the fossil record[4]. In contrast to pterobranchs, the other group of hemichordates, called enteropneusts, are solitary animals that range in length from less than a millimetre[5] to more than 2 metres (Fig. 1). Known as acorn worms, enteropneust adults burrow in soft sediments.

Simakov and colleagues present the genome sequences of two enteropneusts — *Saccoglossus kowalevskii* and *Ptychodera flava*. The authors used these sequences, together with additional DNA sequence data on pterobranchs and several other animals, to build a phylogenetic tree that finds pterobranchs and enteropneusts to be sister groups (Fig. 2). This finding is in agreement with another analysis[6] in rejecting the previously suggested placement of pterobranchs within enteropneusts.

CASEY W. DUNN

Figure 1 | Pharyngeal gill slits. Enteropneusts, better known as acorn worms, use internal gill slits in the pharynx region of their trunk to move water through their mouth to obtain oxygen and, in some species, for filter feeding. The gill slits connect to external gill pores. This specimen is several centimetres long.

As interesting as hemichordates are in their own right[2], much of the motivation for taking a closer look at them comes from a desire to understand their relatives. This is because hemichordates fall within Deuterostomia, the group of animals that also includes echinoderms (radially symmetrical organisms such as sea stars and sea urchins) and chordates. Chordates are of particular interest because they include humans and our vertebrate kin. Although many chordate genome sequences are available, there are few genome resources for other deuterostomes. A draft genome for a sea urchin is available[7], but until now there were no published genomes for hemichordates.

The most recent common ancestor of deuterostomes lived more than 500 million years ago, and there is great diversity in the anatomy of adults of this group. However, many features of deuterostome embryology, including the formation of the anus from the blastopore and the creation of coelomic cavities by pinching off from the gut, are highly evolutionarily conserved. The main finding of Simakov and colleagues' study is that deuterostome genomes, like their embryology, show extensive conservation across great evolutionary timescales. The hemichordate sequences share many features with other deuterostome genomes, including gene composition, exon–intron structure and small- and large-scale gene order. This means that many well-characterized features of chordate genomes are not chordate-specific, but arose earlier in animal evolution.

One of the most conspicuous deuterostome-specific traits is the pharyngeal gill slits. These openings allow water to pass through the mouth without entering the digestive tract, and they are involved in feeding and respiration in these animals. Gill slits arose in the stem lineage that gave rise to deuterostomes, and are not found in non-deuterostome animals, nor in echinoderms, in which they were secondarily lost (Fig. 2). A detailed understanding of the evolutionary origin of this feature is key

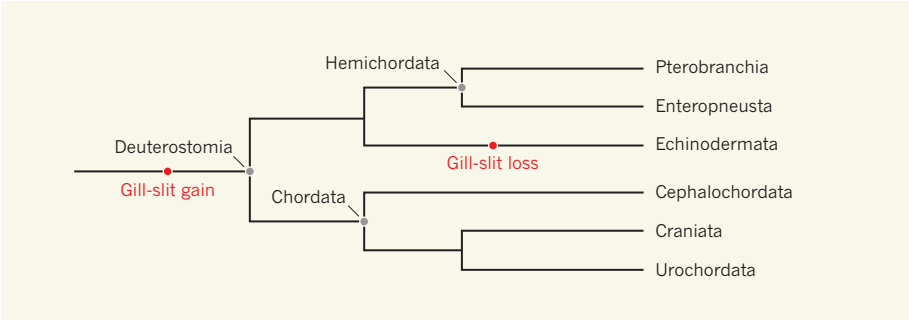


Figure 2 | Deuterostome relationships. The deuterostome group can be divided into three clades: chordates (Cephalochordata, or lancelets; Craniata, which includes all vertebrates; and Urochordata, such as sea squirts); echinoderms (sea stars, sea urchins and relatives); and hemichordates (pterobranchs and enteropneusts). Simakov *et al.*[1] present the first hemichordate genome sequences, from two enteropneust species. The authors' analyses provide new detail on evolutionarily conserved genes that play a part in the development of gill slits. These structures arose along the deuterostome stem, were lost in echinoderms and are reduced in the adults of some chordates (including humans).

to understanding deuterostome, and therefore our own, biology. Simakov *et al.* examine a conserved cluster of six genes that is found only in deuterostomes, and that includes genes known to be involved in patterning gill slits in other deuterostome species. In keeping with the other conserved genome features that they identified, the authors find that these genes are also expressed in the pharyngeal-gill structure of hemichordates.

Simakov and colleagues recognize that there is no 'typical' representative of any animal group; they sequenced two full hemichordate genomes, and collected less-detailed sequence data from a variety of other species to put these genomes in a richer evolutionary context. However, the authors' study still faces the same challenge as all genome investigations — we are far from understanding which evolutionary changes in genomes underlie which evolutionary changes in traits, including development, anatomy and functional biology[8].

There are several reasons for this. First, many evolutionary genome changes are neutral[9] — they have no impact on traits or fitness. This means that we should not assume that any particular genome change affects any traits. Second, on any given phylogenetic branch there will be many changes in both traits and genomes, and there are many possible functional implications for any particular genome change. Third, genome function itself evolves, so the same genome features do not necessarily relate to the same traits in different species.

Our current coarse perspective on genome evolution will improve as more genomes are sequenced, and as functional genomic tools are developed that can be applied to any organism, not just those that can be grown in the laboratory. The conservation of so many features across deuterostome genomes, which is brought into sharp focus with Simakov and colleagues' addition of hemichordate genome sequences, reinforces the fact that radical morphological changes are not necessarily related to radical changes in genomes. This fact will shape the search for which of the variable features of deuterostome genomes are responsible for the great diversity we see across the group. ■

Casey W. Dunn *is in the Department of Ecology and Evolutionary Biology, Brown University, Providence, Rhode Island 02912, USA.*
e-mail: casey_dunn@brown.edu

1. Simakov, O. *et al. Nature* **527,** 459–465 (2015).
2. Kaul-Strehlow, S. & Röttinger, E. in *Evolutionary Developmental Biology of Invertebrates 6: Deuterostomia* (ed. Wanninger, A.) 59–89; http://dx.doi.org/10.1007/978-3-7091-1856-6_2 (Springer, 2015).
3. Lester, S. M. *Mar. Biol.* **85,** 263–268 (1985).
4. Mitchell, C. E., Melchin, M. J., Cameron, C. B. & Maletz, J. *Lethaia* **46,** 34–56 (2013).
5. Worsaae, K., Sterrer, W., Kaul-Strehlow, S., Hay-Schmidt, A. & Giribet, G. *PLoS ONE* **7,** e48529 (2012).
6. Cannon, J. T. *et al. Curr. Biol.* **24,** 2827–2832 (2014).
7. Sea Urchin Genome Sequencing Consortium *et al. Science* **314,** 941–952 (2006).
8. Dunn, C. W. & Ryan, J. F. *Curr. Opin. Genet. Dev.* **35,** 25–32 (2015).
9. Lynch, M. *The Origins of Genome Architecture* (Sinauer, 2007).

This article was published online on 18 November 2015.

CIRCADIAN CLOCKS

A receptor for subtle temperature changes

The protein IR25a is best known for its role as an odour receptor in flies, but an analysis reveals that it also acts to synchronize the circadian clock by sensing small temperature fluctuations. SEE LETTER P.516

FRANÇOIS ROUYER
& ABHISHEK CHATTERJEE

Our body's circadian clocks sense the environmental changes that occur over 24 hours, allowing us to adapt our physiology and behaviour to day–night cycles. Light and temperature have by far the greatest influence on the clock that drives rest–activity rhythms, but how the neurons of this clock synchronize to temperature in the brain remains largely unknown. On page 516 of this issue, Chen *et al.*[1] identify a receptor protein in mechanosensory organs in flies that acts as a specialized temperature sensor, synchronizing the circadian clock with low-amplitude temperature cycles.

Small daily fluctuations of just 1–2 °C are enough to synchronize the fly brain's circadian clock with temperature[2] (a process known as temperature entrainment). Experiments using cultures of different fly body parts have revealed that most organs can entrain their clocks with temperature cycles. The exception is the brain, which must thus rely on external sensors[3]. Expression of the *nocte* gene is required both for the normal development of mechanosensory structures called chordotonal organs and for temperature entrainment of the brain clock[3]. This suggests that chordotonal organs, which are present in the antennae and body parts such as legs and wing hinges, are the external sensors. Although the antennae are the major temperature-sensing organ[4,5], they are not essential for entrainment, indicating that the body chordotonal organs can do the job.

To analyse the role of the body chordotonal organs in temperature entrainment of the brain's rest–activity clock, Chen *et al.* looked for proteins that interact with the Nocte protein. Among the putative Nocte-binding partners that they identified was the protein Ionotropic Receptor 25a (IR25a). IR25a is part of the IR family, members of which are found in sensory organs and are involved in detecting chemicals[6]. So far, IR25a has been best known for its role as a component of a multimeric odour receptor in antennae[7]. As expected for a Nocte partner, the authors found that IR25a

was present in the sensory neurons of chordotonal organs.

Chen and colleagues investigated whether flies lacking IR25a could synchronize their rest–activity rhythms to temperature changes using several different entrainment protocols. For example, they entrained flies using regular 12-hour intervals of light and dark, and then kept the entrained animals in the dark while applying a new regime of large-amplitude temperature cycles (between 16 and 25 °C or 20 and 29 °C), 7 hours ahead of the old light–dark cycle. Both wild-type flies and *IR25a* mutants quickly reset their clock, becoming most active at the end of the warm phase, whereas *nocte* mutants did not. However, when the new regime involved low-amplitude temperature fluctuations (18–20 °C, 21–23 °C and 25–27 °C), the *IR25a* mutants failed to adapt (Fig. 1). IR25a is thus required for synchronizing the circadian system with low-amplitude temperature changes, apparently independently of the temperature range. The authors also provide data to show that IR25a acts in the body chordotonal organs, rather than the antennae.

If IR25a acts as a temperature sensor in sensory neurons upstream of the rest–activity clock, its loss should prevent synchronization with temperature not only of behaviour, but also of the oscillations of the proteins that make up the clock itself. The rest–activity clock involves protein oscillations in about 150 neurons, from around 6 subsets. Chen *et al.* observed complex defects in clock-protein oscillations during shallow temperature cycles in *IR25a* mutants, whereas oscillations were normal during light–dark cycles.

Most clock-neuron subsets were affected, but the authors observed intriguing differences when the temperature cycles were applied in constant light or darkness. Notably, some of the proteins in some groups, such as ventral lateral neurons (LNvs), stopped cycling in darkness, whereas others, such as DN2 dorsal neurons, stopped in light. How temperature information travels in the clock network remains to be discovered, but the authors revealed a key role for DN groups — blocking their neuronal output with tetanus toxin prevented behavioural synchronization to shallow temperature cycles. Interestingly, this finding sits well with previous observations that DN2 neurons are involved in the temperature entrainment of the larval clock[8], and temperature preference rhythms in the adult[9].

Finally, Chen *et al.* demonstrated the role of IR25a in neuronal responses to small temperature changes. Whereas sensory neurons in the legs of both wild-type and *IR25a*-mutant flies responded to movements of the leg, only wild-type neurons were activated by temperature changes. Moreover, when IR25a was misexpressed in a population of large LNv clock neurons, their activity increased in response to small temperature fluctuations. Thus, the presence of IR25A is sufficient to induce neuronal responses to temperature, even in the absence of its partners in the multimeric olfactory receptor.

Chen and colleagues' study deftly demonstrates the temperature sensitivity of the rest–activity clock, and suggests that the chordotonal organs are key players in the clock's temperature entrainment. However, the inability of *nocte* mutants to respond to large temperature cycles indicates that the chordotonal organs also mediate the response of circadian rhythms to larger temperature changes. Whether IR25a might play a part here is unclear, but the process clearly involves other temperature sensors as well.

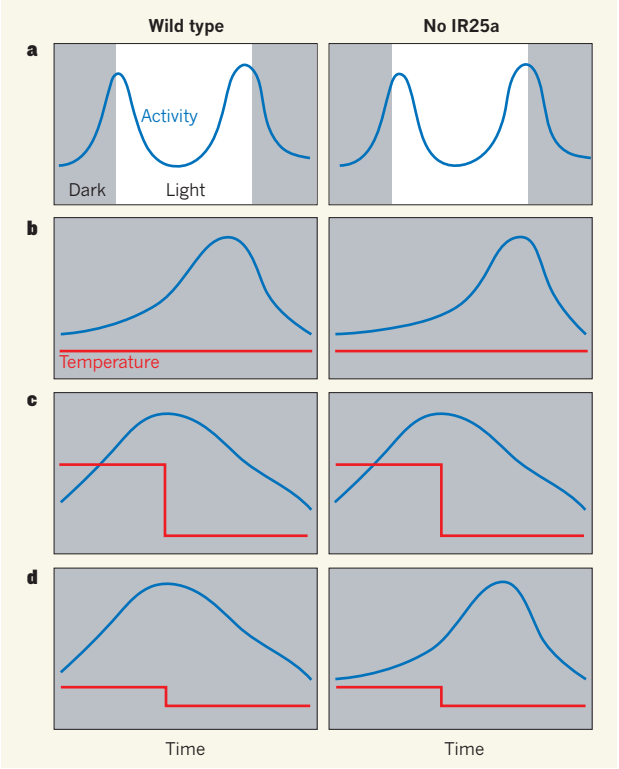


Figure 1 | Telling time with temperature. Chen *et al.*[1] investigated whether flies lacking the protein IR25a can respond to temperature changes to reset the circadian clock in their brain that controls rest–activity rhythms. **a**, The authors synchronized the flies' clock to 12-hour light–dark cycles, in which activity peaked twice during light hours. The flies were then kept in constant darkness and exposed to different temperature cycles. **b**, At constant temperature, the insects' activity changed, but was still synchronized to the original light–dark cycle. **c**, When exposed to large temperature cycles (variations of 9 °C) that fluctuated 7 hours ahead of the light–dark cycle, the clock was reset in both types of fly, so that the activity peak occurred at the end of the warm period. **d**, When exposed to low-amplitude fluctuations (2 °C), flies lacking IR25a could not respond, demonstrating that IR25a mediates the clock's synchronization with small temperature changes.

One candidate is the ion-channel protein Pyrexia, which is found in the chordotonal organs and seems to be needed for entrainment in low-temperature conditions[10]. Another channel, TrpA1, has been implicated in the control of the fly clock by temperature[11]. However, a consensus is emerging for a role for TrpA1 in the temperature-dependent regulation of afternoon siestas, rather than in entrainment[12–14].

The neuronal circuits that pass temperature information from the chordotonal organs to the brain also remain unknown. The complex changes in protein oscillations caused by loss of IR25a point to a role for several groups of neurons. But whether different temperatures or amplitudes of change target specific subsets of clock neuron is an open question. In addition, Chen and colleagues' work suggests that the clock network responds differently to temperature in constant light or darkness, suggesting that light inputs strongly affect temperature entrainment.

Although light can entrain the brain clock through both internal and external sensors, temperature entrainment seems to rely on external sensors. The benefit of preventing the brain clock from directly sensing temperature is unclear. In mammals, the rest–activity clock, which is located in a brain region called the suprachiasmatic nuclei, also controls body-temperature rhythms and uses them to synchronize peripheral clocks[15,16] — it therefore makes sense that these nuclei contain temperature-resistant neuronal networks. But flies do not regulate their body temperature, so why is their brain clock temperature-resistant? Perhaps this organization prevents any overly strong effects of temperature in favour of light, whose daily oscillation might be more reliable. Perhaps it mediates a balance between light and temperature, allowing one entrainment circuit to influence the other. Deciphering how the two sensory modalities are integrated by the clock neuronal network will be an exciting challenge for the next few years. ■

François Rouyer *and* **Abhishek Chatterjee**
are at the Institut des Neurosciences Paris-Saclay, Université Paris-Sud, CNRS, Université Paris-Saclay, 91190 Gif-sur-Yvette, France.
e-mail : francois.rouyer@inaf.cnrs-gif.fr

1. Chen, C. *et al. Nature* **527,** 516–520 (2015).
2. Wheeler, D. A., Hamblen-Coyle, M. J., Dushay, M. S. & Hall, J. C. *J. Biol. Rhythms* **8,** 67–94 (1993).
3. Sehadova, H. *et al. Neuron* **64,** 251–266 (2009).

4. Frank, D. D., Jouandet, G. C., Kearney, P. J., Macpherson, L. J. & Gallio, M. *Nature* **519,** 358–361 (2015).
5. Liu, W. W., Mazor, O. & Wilson, R. I. *Nature* **519,** 353–357 (2015).
6. Benton, R., Vannice, K. S., Gomez-Diaz, C. & Vosshall, L. B. *Cell* **136,** 149–162 (2009).
7. Abuin, L. *et al. Neuron* **69,** 44–60 (2011)
8. Picot, M., Klarsfeld, A., Chélot, E., Malpel, S. & Rouyer, F. *J. Neurosci.* **29,** 8312–8320 (2009).
9. Kaneko, H. *et al. Curr. Biol.* **22,** 1851–1857 (2012).
10. Wolfgang, W., Simoni, A., Gentile, C. & Stanewsky, R. *Proc. R. Soc. B* **280,** 20130959 (2013).
11. Lee, Y. & Montell, C. *J. Neurosci.* **33,** 6716–6725 (2013).
12. Das, A., Holmes, T. C. & Sheeba, V. *PLoS ONE* **10,** e0134213 (2015).
13. Green, E. W. *et al. Proc. Natl Acad. Sci. USA* **112,** 8702–8707 (2015).
14. Roessingh, S., Wolfgang, W. & Stanewsky, R. *J. Biol. Rhythms* http://dx.doi.org/10.1177/0748730415605633 (2015).
15. Brown, S. A., Zumbrunn, G., Fleury-Olela, F., Preitner, N. & Schibler, U. *Curr. Biol.* **12,** 1574–1583 (2002).
16. Buhr, E. D., Yoo, S. H. & Takahashi, J. S. *Science* **330,** 379–385 (2010).

This article was published online on 18 November 2015.

IMAGING TECHNIQUES

Super-resolution ultrasound

By infusing blood vessels with gas-filled microbubbles and using rapid ultrasound imaging to detect the bubbles, super-resolution imaging of an entire vessel system has been achieved in a rat brain. SEE LETTER P.499

BEN COX & PAUL BEARD

Ultrasound imaging is used in hospitals throughout the world as a safe, non-invasive and relatively inexpensive way to visualize a patient's internal tissues in real time. The quality of ultrasound images has been improving steadily since the 1970s, owing to advances in hardware and image-forming algorithms. But, like all wave-based imaging techniques, ultrasound faces a fundamental limit because of the way that waves spread out (diffract) as they travel — two objects are distinguishable from one another only if they are more than half a wavelength apart. On page 499 of this issue, Errico *et al.*[1] overcome this limit to produce super-resolution images of the microvasculature in the brain of a live rat.

If the resolution limit of ultrasound imaging is related to wavelength, why not just use sound with a shorter wavelength? Although this approach is useful to some extent, the absorption of ultrasound waves increases strongly as the wavelength decreases; therefore, using shorter wavelengths limits the depth to which tissue can be imaged before the reflected waves are attenuated too much to be detected. As a result, the resolution limit in clinical ultrasound imaging is at best hundreds of micrometres. To take useful images at depth, it is therefore necessary to bypass the half-wavelength limit.

The same fundamental resolution limit is found in light microscopy. But the development of several super-resolution techniques, such as photoactivated localization microscopy (PALM), has enabled researchers to achieve nanoscale resolution — breakthroughs for which the 2014 Nobel Prize in Chemistry was awarded.

PALM achieves super-resolution imaging in three steps. The first step is to image light-activated fluorescent molecules that act as tiny, randomly distributed pinpricks of light. The use of low light intensities and the fact that the molecules' activation is inherently random ensures that only a sparse subset is turned on at any one time. Thus, these point-like light sources are separated by more than half a wavelength, so the image of each one (a blurred spot called the point spread function) does not overlap with that of its neighbours.

The second step is to determine the exact position of each point-like source by finding the centre of the point spread function. This is possible for well-separated sources, because the shape of the point spread function can be known in advance. The final step is to repeat the illumination and detection steps many times. A different set of separated point-like sources is detected each time, until a sufficient density of source points has been obtained. By marking the positions of all of these point sources on a single meta-image, a super-resolved picture can be built up. The spatial resolution in this image can exceed the diffraction limit, because it is determined by the accuracy with which the position of each source can be estimated.

Could a similar approach be used to achieve super-resolution ultrasound imaging? The first challenge is to identify potential point sources (point scatterers in the case of ultrasound). Because small blood vessels are poor at reflecting sound waves, they can be hard to see using ultrasound, and gas-filled microbubbles, which reflect sound well, have long been used as contrast agents to enhance vessel visibility. Microbubbles are strong scatterers of sound, and so are good candidates as point sources. But to be useful in this context, there must be some way to identify them separately in ultrasound images.

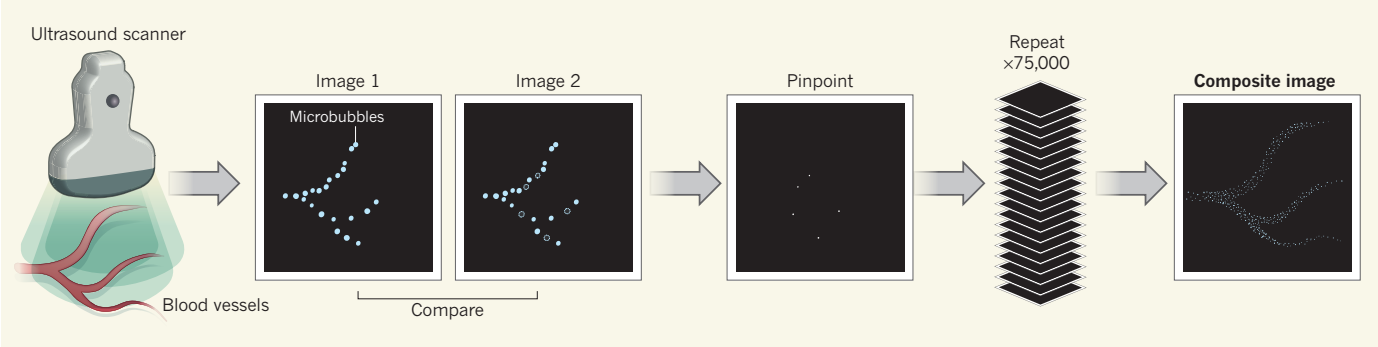


Figure 1 | Vessels visualized. Errico *et al.*[1] obtained ultrasound images of rat blood vessels infused with gas-filled microbubbles, which reflect ultrasound waves. By taking images rapidly (at 500 frames per second) and generating difference data to compare sequential images, they were able to pinpoint the locations of the few, well-separated microbubbles that degraded between each image (only a small number of microbubbles are shown here for simplicity, and are not to scale). By repeating this process over many frames, a composite image was built up that revealed the locations of many thousands of microbubbles, making super-resolution images possible in around 150 seconds.

In 2013, researchers achieved super-resolution ultrasound imaging by using a sufficiently dilute solution of microbubbles to achieve the necessary separation[2]. Earlier this year, the same group used this approach to obtain super-resolution images of the microvasculature in a mouse ear to a depth of more than one centimetre[3]. They also tracked the microbubbles, to estimate blood-flow velocity. However, their system acquired images at the low rate of 25 frames per second, which meant that hour-long imaging times were needed to achieve super-resolution.

Errico *et al.* used a different approach that dispensed with the need for a dilute microbubble solution. By using a high-frame-rate imaging system (500 frames per second), they were able to detect the waves scattered from individual microbubbles by comparing sequential images. Signals from bubbles that have disintegrated or moved significantly in the time between frames can be detected in the data, and — as long as these changes are sparse enough to be spatially separated from one another — the positions and velocities of these microbubbles can be accurately determined[4,5] (Fig. 1). The authors compared 75,000 images taken over about 150 seconds to build up super-resolution images of the vasculature in the cortical region of rat brains, through both intact rat skulls and skulls that had been thinned to reduce the acoustic attenuation.

Could this technique be translated to a clinical setting? At the ultrasound wavelengths used by Errico and colleagues, overcoming the attenuating effect of the thick human skull will present a considerable challenge. The authors point out that it might be possible to circumvent this problem by using longer wavelengths, which are less severely attenuated. Nonetheless, imaging of less-challenging targets that do not require the ultrasound waves to pass through thick bone should be readily achievable. One disadvantage of the new approach compared with conventional ultrasound imaging is the need to administer a contrast agent; this requires an intravenous cannula and can increase clinical scanning time.

Super-resolution ultrasound imaging of microvasculature is an exciting prospect. The technique has the potential to substantially advance the study of normal blood-vessel function, as well as disease. Moreover, it might enable doctors to readily identify microvessel-related disorders, such as tumour-related vessel growth and microvascular abnormalities in deep abdominal organs such as the kidneys, and to assess cardiovascular disease. ■

Ben Cox *and* **Paul Beard** *are in the Department of Medical Physics and Biomedical Engineering, University College London, London WC1E 6BT, UK.*
e-mail: b.cox@ucl.ac.uk

1. Errico, C. *et al. Nature* **527,** 499–502 (2015).
2. Viessmann, O. M., Eckersley, R. J., Christensen-Jeffries, K., Tang, M. X. & Dunsby, C. *Phys. Med. Biol.* **58,** 6447–6458 (2013).
3. Christensen-Jeffries, K., Browning, R. J., Tang, M. X., Dunsby, C. & Eckersley, R. J. *IEEE Trans. Med. Imag.* **34,** 433–440 (2015).
4. Couture, O., Besson, B., Montaldo, G., Fink, M. & Tanter, M. *Proc. IEEE Int. Ultrasonics Symp.* 1285–1287 (2011).
5. Desailly, Y., Couture, O., Fink, M. & Tanter, M. *Appl. Phys. Lett.* **103,** 174107 (2013).

CELL FATE

Transition loses its invasive edge

Two studies provide evidence that epithelial tumour cells do not need to transition to a mesenchymal–cell state to form metastases, but that this process does contribute to drug resistance. **SEE ARTICLE P.472 & LETTER P.525**

SHYAMALA MAHESWARAN & DANIEL A. HABER

Cancer often becomes lethal only when cells from the primary tumour disseminate to another organ. The early steps of this highly complex process, called metastasis, have been thought[1] to rely on non-motile epithelial tumour cells acquiring characteristics of mesenchymal cells, which are more migratory. This change is known as the epithelial-to-mesenchymal transition (EMT). The migrating cancer cells then undergo a reverse mesenchymal-to-epithelial transition when they seed a secondary tumour[2]. Metastases therefore display the same epithelial-cell predominance as primary cancers, leaving no evidence of their transient mesenchymal state. Now, two papers in this issue (Fischer *et al.*[3] and Zheng *et al.*[4]) present data that challenge the role of EMT as a crucial effector of cancer metastasis.

In terms of cancer-cell characteristics, the epithelial lineage is associated with increased proliferation, whereas the mesenchymal lineage is linked to enhanced avoidance of anoikis (a form of death that occurs when cells detach from their normal tissue matrix) and of drug-induced death. Current understanding of EMT-induced metastasis is derived from *in vitro* data and mouse models of human cancers. But the inherent plasticity of the process and the limited clinical evidence supporting the occurrence of EMT in tumour specimens[2,5] have led to scepticism about EMT being the predominant mechanism governing the early steps of metastasis. EMT is, however, emerging as one of numerous mechanisms conferring resistance to various cancer therapies[6].

Using mouse models of mammary tumours, Fischer and co-workers (page 472) surveyed the fate of epithelial tumour cells transitioning to a mesenchymal state, from the cells' inception and dissemination through the bloodstream to their exit from blood vessels and metastatic growth. To do this, the authors monitored the expression of green fluorescent protein (GFP) as a proxy for the expression of the genes that encode fibroblast-specific protein 1 (FSP1) or vimentin, which are triggered when epithelial tumour cells switch to a mesenchymal state. The green fluorescence persists in the progeny of these cells well after they revert to an epithelial fate.

In two mouse models, the authors found evidence for EMT in a minor fraction of cells in the primary tumour and in a subset of circulating tumour cells. However, the vast majority of metastatic tumours were not derived from the mesenchymal-switched cells expressing GFP, but from disseminating epithelial cells (Fig. 1). The researchers further show that inhibiting expression of the genes *ZEB1* and *ZEB2*, which locks tumour cells in the epithelial state, did not impair metastasis of the mouse mammary tumours to the lung.

Zheng *et al.* (page 525) reached a similar conclusion using tissue-specific deletion of the EMT-inducing transcription factors Snail or Twist to assess the consequences of EMT in a mouse model of pancreatic cancer. They found that loss of either Snail or Twist in the pancreatic epithelium does not affect tumour formation or overall survival, but it does suppress EMT in the primary tumour. Despite the lower frequency of cells expressing mesenchymal marker proteins in these tumours compared with those in mice in which Snail and Twist are expressed at normal levels, there were a similar number of metastases in the liver, lungs and spleen of these mice.

Although they question the role of EMT in metastatic dissemination, both research groups go on to conclude that EMT contributes to drug resistance. In Fischer and colleagues' mammary-tumour model, treatment with the drug cyclophosphamide led to enhanced survival and proliferation of

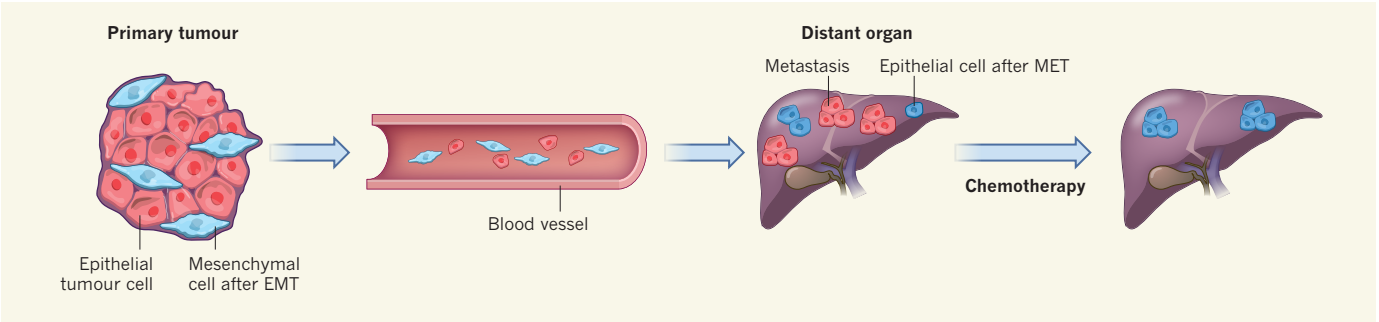


Figure 1 | Metastatic potential. A small fraction of epithelial cells in a solid tumour acquires mesenchymal-cell characteristics during tumour progression, through a process known as epithelial-to-mesenchymal transition (EMT). Both epithelial and mesenchymal cancer cells can invade the bloodstream and exit it at distant sites, where the mesenchymal cells undergo a reverse mesenchymal-to-epithelial (MET) transition. Contrary to previous opinion, Fischer *et al.*[3] and Zheng *et al.*[4] find that the majority of metastatic tumours at secondary sites are initiated by epithelial cells from the primary tumour, and not by cells that have undergone EMT and subsequent MET. However, both research groups also show that such transitioned cells are more resistant to chemotherapeutic drugs than are untransitioned epithelial cells and emerge as the dominant metastatic population following treatment.

EMT-switched tumour cells that had reached the lung. Analysis of the genes being expressed in these cells correlated this resistance with increased expression of genes encoding drug-metabolizing enzymes and drug-transporter proteins. These findings were mirrored in Zheng and colleagues' Snail- or Twist-deleted pancreatic cancers, in which tumour cells with epithelial characteristics expressed higher levels of nucleoside-transporter proteins than did mesenchymal cells, potentially rendering the epithelial cells more sensitive to the chemotherapeutic drug gemcitabine.

These findings challenge the prevailing hypothesis that EMT is a key element in the metastatic dissemination of epithelial cancers, and they point to a distinct role of this cell-fate transition in enhancing cancer-cell survival during drug treatment. How can these data be reconciled with compelling previous reports on the role of EMT in metastasis? The earlier work studied the effects of EMT induced by the growth factor TGFβ or by overexpression of transcriptional regulators such as Snail or Twist. Such approaches might not capture the physiological process that occurs spontaneously in cancer cells as accurately as do the methods used in the current studies. It is probable that the induction of EMT that occurs during the natural progression of a cancer may be more subtle than the full EMT switch that is induced by the expression of powerful regulators and that has been associated with high levels of metastasis.

However, the limitations of the models used by Fischer *et al.* and Zheng *et al.* need to be considered before EMT-mediated tumour invasion can be dismissed outright. EMT is orchestrated by complex circuitry involving multiple signalling molecules and transcription factors. Tracing switched cells on the basis of expression of a single gene may therefore not fully capture these complicated features. Similarly, Snail and Twist function redundantly in many settings[2,7,8], and inactivation of both (and potentially of other transcriptional regulators) simultaneously, rather than individually, may be required to abrogate EMT. Furthermore, cancer is a highly variable disease, and its full complexity cannot be completely captured in mouse models that are driven by expression of a few cancer-initiating genes. Nonetheless, the conclusions reached by the two studies warrant a re-evaluation of the role of EMT in cancer progression. Alternative ways in which epithelial cells could enter the bloodstream without acquiring mesenchymal properties, such as collective epithelial-cell migration[9] or tumour fragmentation[10], are worth investigating.

The postulated role of EMT in mediating cancer-cell survival is reinforced by the two latest studies. Indeed, EMT has been linked to drug susceptibility of cancer cells, as well as to their entrance into a non-proliferative state in which they have stem-cell-like properties[11]. Understanding the many cellular pathways that together determine these cell fates, and how these pathways are modulated, is likely to provide fertile ground for drug discovery and for new therapeutic strategies. ■

Shyamala Maheswaran *and* **Daniel A. Haber**
are at the Massachusetts General Hospital Cancer Center, Harvard Medical School, Boston, Massachusetts 02114, USA.
e-mails: maheswaran@helix.mgh.harvard.edu; dhaber@mgh.harvard.edu

1. Thiery, J. P., Acloque, H., Huang, R. Y. & Nieto, M. A. *Cell* **139,** 871–890 (2009).
2. Ye, X. & Weinberg, R. A. *Trends Cell Biol.* **25,** 675–686 (2015).
3. Fischer, K. R. *et al. Nature* **527,** 472–476 (2015).
4. Zheng, X. *et al. Nature* **527,** 525–530 (2015).
5. Yu, M. *et al. Science* **339,** 580–584 (2013).
6. Singh, A. & Settleman, J. E *Oncogene* **29,** 4741–4751 (2010).
7. Peinado, H., Olmeda, D. & Cano, A. *Nature Rev. Cancer* **7,** 415–428 (2007).
8. Puisieux, A., Brabletz, T. & Caramel, J. *Nature Cell Biol.* **16,** 488–494 (2014).
9. Clark, A. G. & Vignjevic, D. M. *Curr. Opin. Cell Biol.* **36,** 13–22 (2015).
10. Aceto, N. *et al. Cell* **158,** 1110–1122 (2014).
11. Rumman, M., Dhawan, J. & Kassem, M. *Stem Cells* **33,** 2903–2912 (2015).

This article was published online on 11 November 2015.

EBOLA

Hidden reservoirs

West Africa's Ebola epidemic continues to reveal surprises. Although the animal species that originally passed the virus to people remains a mystery, a virus reservoir and persistent disease have been identified in some human survivors.

JONATHAN L. HEENEY

Animals are reservoirs for many pathogens that occasionally jump species and infect humans. In December 2013 in the forests of Guinea, a two-year-old boy became infected with the Zaire strain of Ebola virus from an unidentified animal source[1]. This event triggered the largest and longest human epidemic of Ebola viral infection in recorded history. Across several countries in West Africa, over 28,000 people were infected and more than 11,000 died. This fatality rate of less than 50% was lower than in most previous outbreaks, and it left more than 16,000 survivors[2]. Studies of these survivors are changing our understanding of Ebola virus infection and raising concern for the long-term well-being of these individuals and their communities. Writing in the *New England Journal of Medicine*, Deen *et al.*[3] reveal that Ebola virus RNA can persist in the semen of men for months

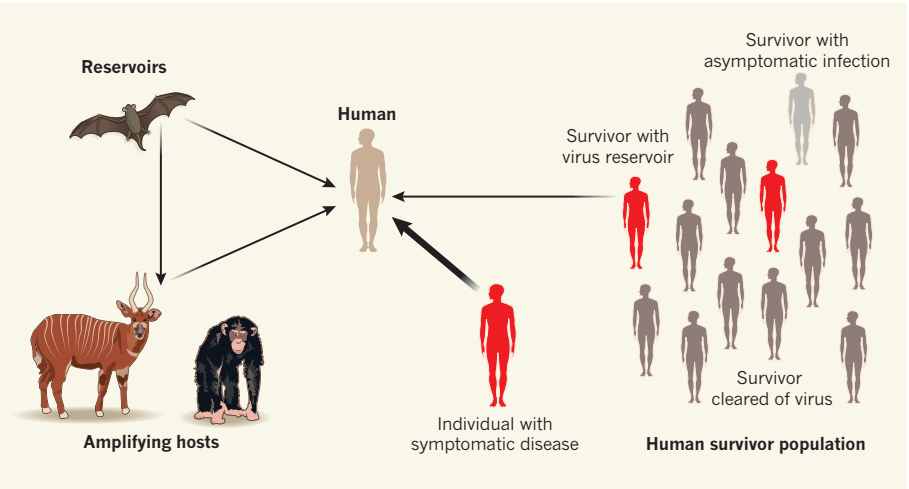


Figure 1 | Ebola infection dynamics in animals and humans. Ebola virus has been identified in several animal species, including bats, chimpanzees and forest antelopes. Transmission to humans can occur directly from reservoir species, in which the virus may persist without causing active infection, or from amplifying host species, in which the virus replicates to high levels, often causing illness and death. Most infected people develop acute Ebola virus disease and are highly infectious, although some individuals survive exposure and infection without developing symptoms. There is also growing evidence[3,4] that the virus can persist in the central nervous system and reproductive organs of some survivors of the disease, with the possibility that these survivors could infect others months after resolution of their acute symptoms[4].

after their recovery from the disease, and Mate *et al.*[4] demonstrate that such persistence can be the source of new infections through sexual transmission.

Deen and colleagues obtained semen samples from 93 Sierra Leonian men who had survived Ebola virus disease (EVD) at various intervals after the onset of their disease. Although the authors find that the proportion of men whose semen contained Ebola virus RNA waned with time, the viral genomes persisted for as long as 7–9 months after recovery. Mate and colleagues provide convincing evidence that a female patient in Liberia, who subsequently died, had contracted Ebola virus through unprotected vaginal intercourse with her male partner, who had survived EVD.

These observations support similar findings in previous epidemics of filoviruses, the virus family to which Ebola belongs. There have been reports[5,6] of the persistence of Marburg virus in the anterior chamber of the eye and semen of human survivors, and of the persistence of Ebola virus in the semen of men who survived the 1995 outbreak in the Democratic Republic of the Congo[7]. This has obvious implications for sexual partners.

The fact that Ebola virus is found at high levels in placental tissues also suggests that transmission could occur from pregnant women who survive EVD to their babies, although pregnant women who become infected usually abort the fetus before term[8]. Mother-to-child transmission by breastfeeding in survivors of Marburg virus has been reported[9], and the potential for transmission through breast milk has also been suggested for Ebola[10].

Although the relative risk of virus transmission from survivors is low compared with transmission from patients with acute EVD, a single case of new infection is sufficient to trigger an epidemic (Fig. 1). Thus, there is a strong need for rigorous assessment of the tissue reservoirs of Ebola in human survivors and the associated public-health risks. Follow-up health care should be combined with compassionate education of survivors and their communities by qualified and knowledgeable personnel, including advice on condom use.

Another lesson to emerge from this epidemic is that some survivors experience symptoms after their recovery from the main disease episode, suggesting that viral persistence in certain compartments of the body is more serious in some survivors than previously recognized. Reported symptoms include blurred vision, pain behind the eyes, hearing deficits, painful swallowing, joint pain, fever, memory loss and difficulty in sleeping[11,12]. The rehospitalization of a British nurse who developed neurological complications more than 9 months after surviving acute EVD[13] is a chilling indication that the virus can persist in the central nervous system and be triggered to reactivate or to escape immune surveillance, or both. Fortunately, diagnosis and successful clinical intervention were possible for the nurse in Britain, but this situation is unlikely in most communities in West Africa.

The existence of a reservoir state in human Ebola survivors is now beyond debate. But we do not know how long viable virus can persist in these tissue reservoirs, nor whether the virus replicates there at low levels or is dormant and then triggered to replicate. Better definition and understanding of the reservoirs and the underlying mechanisms of post-EVD symptoms are needed to inform clinical management and treatment.

For example, studies of survivors may identify features of their immune responses (such as neutralizing-antibody determinants) that correlate with either full viral clearance or the persistence of viral reservoirs. Such correlates may enable survivors to be classified into 'carrier' or 'cleared' subtypes. Potential factors that could predispose survivors to viral re-emergence also need to be taken into account, including genetics, compromised immunity owing to poor health, concurrent infections such as HIV, or use of immunosuppressive drugs. However, Ebola, like other RNA viruses, may be prone to mutational changes, and virus escape from the host's immune response may eventually occur even without predisposing factors.

It is also not clear how, or whether, post-EVD immunity is affected by the stage of treatment or type of therapy given, such as monoclonal antibodies or the antibodies in convalescent plasma. As well as helping to classify survivors, enhanced understanding of viral persistence will help to guide therapeutic choices — treatment with small antiviral molecules, for example, may facilitate full clearance of the virus.

Although we are learning much about Ebola from this epidemic, we have yet to identify the events that caused the virus to jump to the Guinean boy almost two years ago. The consumption of bushmeat has been associated with previous epidemics, and some bushmeat species, such as great apes and forest antelopes, are susceptible to high levels of Ebola-virus replication and die from the infection. They are thus best considered as amplifying hosts, rather than the initial reservoir species (Fig. 1). Prime suspects for the reservoir include several species of bat, although a bat source has not been confirmed for this latest epidemic[14]. Indeed, the animal reservoirs of Ebola may be cloaked by sequestration of the virus in much the same way as its persistence in human survivors, waiting for physiological triggers for transmission to unexposed animals of the same species or to amplifying hosts.

Understanding the triggers of Ebola emergence, the persistence of the virus in humans and the infection dynamics in its animal reservoirs is vital not only for the long-term care of survivors of this epidemic, but also for preventing the next one. ■

Jonathan L. Heeney *is in the Laboratory of Viral Zoonotics, University of Cambridge, Cambridge CB3 OES, UK.*
e-mail: jlh66@cam.ac.uk

1. Baize, S. *et al. N. Engl. J. Med.* **371,** 1418–1425 (2014).
2. World Health Organization. http://apps.who.int/ebola/ebola-situation-reports
3. Deen, G. F. *et al. N. Engl. J. Med.* http://dx.doi.org/10.1056/NEJMoa1511410 (2015).

4. Mate, S. E. *et al. N. Engl. J. Med.* (2015). http://dx.doi.org/10.1056/NEJMoa1509773 (2015).
5. Martini, G. A. *Trans. R. Soc. Trop. Med. Hyg.* **63,** 295–302 (1969).
6. Smith, D. H. *et al. Lancet* **319,** 816–820 (1982).
7. Rodriguez, L. L. *et al. J. Infect. Dis.* **179** (Suppl. 1), S170–S176 (1999).
8. Baggi, F. M. *et al. Eurosurveillance* www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20983 (2014).
9. Borchert, M. *et al. Trop. Med. Int. Health* **7,** 902–906 (2002).
10. Bausch, D. G. *et al. J. Infect. Dis.* **196** (Suppl. 2), S142–S147 (2007).
11. Gulland, A. *Br. Med. J.* **351,** h4336 (2015).
12. Clark, D. V *et al. Lancet. Infect. Dis.* **15,** 905–912 (2015).
13. *Nursing Stand.* **30** (8), 8 (2015).
14. Pigott, D. M. *et al. eLife* http://dx.doi.org/10.7554/eLife.04395 (2014).

PLANETARY SCIENCE

The Moon's tilt for gold

The Moon's current orbit is at odds with theories predicting that its early orbit was in Earth's equatorial plane. Simulations now suggest that its orbit was tilted by gravitational interactions with a few large bodies. SEE LETTER P.492

ROBIN CANUP

Four and a half billion years ago, a giant impact with Earth is thought to have created an Earth-orbiting disk of debris that coalesced to form the Moon. 'Inelastic' collisions between such debris would dissipate energy and remove relative up-and-down motions, so that the Moon that assembled from these collisions would orbit approximately in Earth's equatorial plane. Yet the Moon's current orbit implies that its initial orbit was substantially inclined relative to Earth's Equator[1], a troubling contradiction.

On page 492 of this issue, Pahlevan and Morbidelli[2] identify a compelling and simple solution to this problem — that the Moon's early orbit was gravitationally jostled into a tilted state by close passes of large objects left over from the formation of the inner planets. The existence of a population of these objects could also explain how elements such as iridium, platinum and gold were delivered to Earth's outer layers after the Moon formed[3].

The Earth–Moon pair is a dynamically coupled system. The Moon's gravity raises tides on Earth, most notably in the oceans, and gravitational interactions between these tides and the Moon is causing Earth's rotation to slow as the Moon's orbit expands. Tidal interactions also reduce the tilt of the Moon's orbit relative to a preferred plane. This would have coincided with Earth's equatorial plane when the early Moon was orbiting close to Earth and transitioned to the plane of Earth's orbit around the Sun as the Moon's orbit expanded. In the absence of other effects, the current 5° inclination of the Moon's orbit relative to Earth's orbital plane implies an initial 10° inclination relative to Earth's equatorial plane when the Moon formed[1], 10 times larger than expected according to theory[4].

A seemingly unrelated — until now — set of clues about the conditions soon after the Moon formed emerge from the abundance of precious metals in the Earth. Elements such as platinum and gold are highly siderophile, which means that they have strong chemical affinities for iron. Because Earth formed in a largely molten state, high-density iron would have readily sunk to the planet's centre to form a core, taking highly siderophile elements with it and efficiently removing these from Earth's upper layers. The fact that we find such elements in relatively high abundance in rocks at Earth's surface suggests that they were delivered to the planet after the end of core formation, through a 'late veneer' of material that added about the last 1% of Earth's mass[5].

If Earth's late veneer was delivered by a large number of small impactors, the Moon would have received about 1/20th as many impactors on the basis of its smaller cross section[3]. But lunar siderophile abundances imply that the Moon received much less than that amount. It thus seems probable that Earth's late veneer was delivered by only a few large impactors, each roughly comparable in size to the Moon, because the Moon would have received less than its proportionate share under these circumstances.

Pahlevan and Morbidelli use computational methods (Monte Carlo simulations) to consider the effects of such a population of large, late-accreting background objects on the Moon's early orbit. Their simulations begin with a Moon orbiting in Earth's equatorial plane close to our planet (Fig. 1). With time, the Moon's orbit expands because of tidal interaction with Earth, and is gravitationally perturbed by the background objects until this population is depleted over typically a few tens of millions of years.

Central to the new work is the recognition that each object that ultimately collides with Earth first undergoes many thousands of non-collisional close passes, a portion of which strongly perturb the Moon's orbit. An object approaching the Moon from a random

50 Years Ago

The use of rubber gloves during surgical operations became general about 1900 … The object of an investigation was to obtain an estimation of how frequently wound infection originates from bacteria on the hands of operating staff … Examination of the wounds following 433 'clean' operations, of the 3,125 rubber gloves used in those operations and of the bacterial flora of the hands which had worn 692 damaged gloves, revealed no connexion between the glove damage, the bacterial flora and the wound infections observed.
From *Nature* 27 November 1965

100 Years Ago

The *Times* of November 20 published a rather flamboyant little article, headed "A Surgical Schism." This article said: "Not for half a century at least has the medical world been so sharply divided as it is to-day in regard to the question of the treatment of wounds." Now, it is exactly half a century since Lister … first ventured to treat a compound fracture by plugging the wound with a strip of rag soaked in undiluted and impure German creasote. Pyaemia and septicaemia and erysipelas were ravaging the wards of the old Glasgow Infirmary, and he, relying on Pasteur's work on the "germs of putrefaction," and knowing that creasote was a good "disinfectant," plugged a wound with it. That was the beginning of everything, exactly half a century ago. To-day, there are many methods, but they do not all contradict or exclude each other … We must not imagine a sort of desperate squabble among our military surgeons … The suggestion in the *Times* article that an acute controversy is proceeding upon these matters is unfortunate and misleading.
From *Nature* 25 November 1915

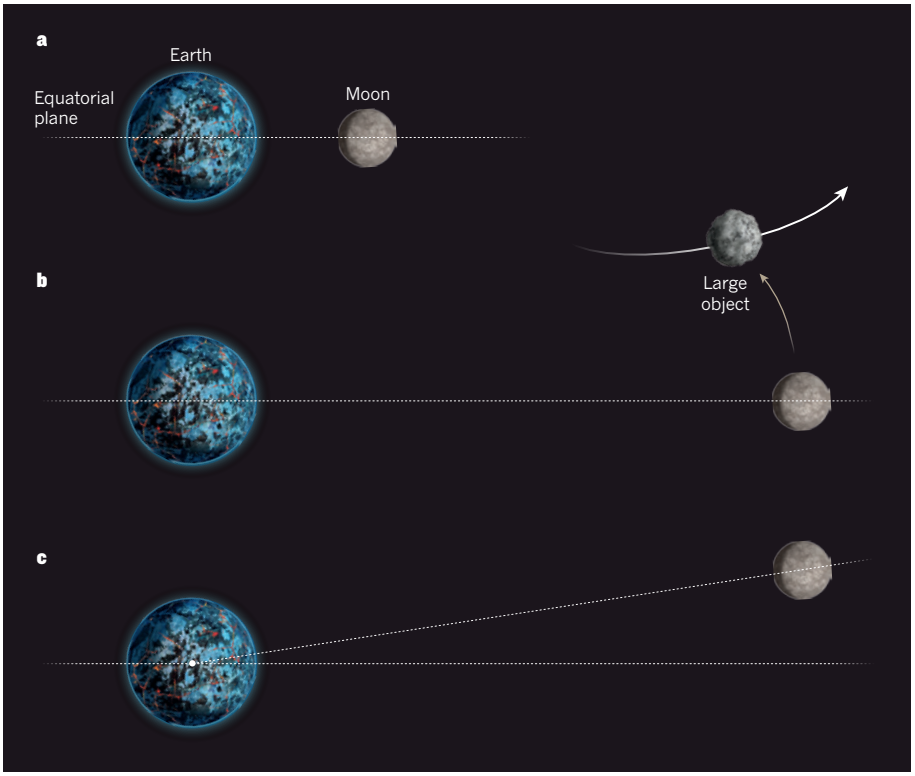


Figure 1 | Collisionless interactions could have altered the early Moon's orbit. **a**, When the Moon first formed, its orbit was approximately in the plane of Earth's Equator. Over time, its orbit then expanded. **b**, Pahlevan and Morbidelli[2] propose that collisionless interactions with large objects passing through the early Solar System would have strongly perturbed the Moon's orbit. **c**, The cumulative effect of such interactions would have tilted the Moon's orbital plane sufficiently to explain the current inclination of the Moon to Earth's orbital plane around the Sun. The Moon's orbital radius in **b** and **c** is not shown to scale.

direction may increase or decrease the Moon's orbital tilt. But just as a series of steps, each equally likely to be forward or backward, causes the standard deviation in the net distance travelled to increase with time, so too does a series of randomly oriented kicks to the lunar orbit lead to a general increase with time in the probability of exciting a minimum tilt.

Pahlevan and Morbidelli's results show a high likelihood that such random scattering events can cumulatively produce the necessary early tilt in the Moon's orbit, as long as the number of objects that deliver the final approximately 1% of Earth's mass is small (fewer than 5) and the rate of early tidal expansion of the Moon's orbit is sufficiently rapid. The rate of early tidal expansion needed is broadly consistent with the average tidal properties inferred for Earth on the basis of the expansion of the Moon's orbit to its current orbital distance. However, the specific values that would have applied to the earliest Earth remain uncertain.

The magnitude of the excited tilt scales roughly linearly with the late mass delivered to Earth. It is not known what fraction of the siderophiles that were concentrated in the cores of such large impactors would have been retained in Earth's upper layers. Improved models of late-veneer impacts should therefore be used to better constrain the late-accreted mass; this would in turn allow a closer approximation of the inclination expected from scattering. Moreover, the new scattering model is most effective if the Moon's inclination has been damped only by tides. If other forms of inclination damping have occurred, then — depending on the timing of this damping — the required initial inclination might increase, and with it the required mass of background objects, perhaps to unrealistically high values.

Previously reported models for the origin of the Moon's inclination rely on more-complex processes involving either a periodic gravitational interaction (gravitational resonance) with the Sun[6] or a resonant interaction between the Moon and its precursor disk[7]. Both require rather narrow sets of conditions for success. The new mechanism is simpler than these models, and the population of late lunar-sized objects that it requires is compellingly consistent with that needed to account for the delivery of Earth's precious metals, a completely independent constraint. Had such a population of objects not existed, the Moon might be orbiting in Earth's orbital plane, with total solar eclipses occurring as a spectacular monthly event. But our jewellery would be much less impressive — made from tin and copper, rather than from platinum and gold. ■

Robin Canup *is in the Planetary Science Directorate, Southwest Research Institute, Boulder, Colorado 80302, USA.*
e-mail: robin@boulder.swri.edu

1. Goldreich, P. *Rev. Geophys.* **4,** 411–439 (1966).
2. Pahlevan, K. & Morbidelli, A. *Nature* **527,** 492–494 (2015).
3. Bottke, W. F., Walker, R. J., Day, J. M. D., Nesvorny, D. & Elkins-Tanton, L. *Science* **330,** 1527–1530 (2010).
4. Ida, S., Canup, R. M. & Stewart, G. R. *Nature* **389,** 353–357 (1997).
5. Walker, R. J. *Chem. Erde* **69,** 101–125 (2009).
6. Touma, J. & Wisdom, J. *Astron. J.* **115,** 1653–1663 (1998).
7. Ward, W. R. & Canup, R. M. *Nature* **403,** 741–743 (2000).

BLINDNESS

Assassins of eyesight

A molecular cascade involving the transcription factor SIX6 and its target gene *p16INK4a* causes the death of neurons that link the eye to the brain. This discovery deepens our understanding of a common form of blindness, glaucoma.

ANDREW D. HUBERMAN & RANA N. EL-DANAF

Vision might feel easy, but an immense number of neurons are required to perform routine visual functions, such as reading, navigating the street or recognizing faces. Tightly lining the back of the eye is a layer of approximately 1 million neurons called retinal ganglion cells (RGCs), which take information encoded by the retina and pass it to the brain[1]. Glaucoma — a disease marked by progressive, irreversible degeneration of RGCs — is a common form of blindness, affecting more than 60 million people worldwide[2]. Although many studies have sought to understand the cellular and molecular basis of glaucoma[3], the mechanisms that drive RGC death in this debilitating disease have remained mysterious. But writing in *Molecular Cell*, Skowronska-Krawczyk *et al.*[4] report that certain glaucoma-associated mutations in humans are linked to a defined molecular pathway that accelerates RGC ageing and death.

A constellation of risk factors has been

associated with glaucoma, one of the greatest of which is age. Like other forms of neurodegeneration, loss of RGCs occurs more often in people over 60, raising questions about whether similar mechanisms might underlie glaucoma and other age-related neurodegenerative disorders such as Alzheimer's disease[5]. There also seems to be a strong genetic component to glaucoma, with certain forms occurring four to five times more frequently in dark-skinned people[6]. Finally, the disease is often thought to be caused by elevated fluid pressure inside the eye. However, abnormally high intraocular pressures are neither 100% predictive of nor a prerequisite for glaucoma, and many people with the disease have normal eye pressures[2]. This broad range of risk factors has led many to speculate that glaucoma is caused by a variety of individual stressors that all increase RGC susceptibility to death. The key questions have therefore become: what are the common molecular pathways that trigger RGC loss, and how could those pathways be manipulated for therapies?

Skowronska-Krawczyk *et al.* analysed genetic-association studies in several human populations to find genes that are commonly mutated in people with primary open-angle glaucoma (the most common form of the disease). One screen picked up *SIX6*, which encodes a transcription factor that helps to shape the eye during embryonic and postnatal development[7]. A mutation called His141, which changes amino-acid residue 141 of the SIX6 protein from asparagine to histidine, confers a risk of glaucoma. The authors performed a careful structural analysis, which revealed that this residue probably lies outside the transcription factor's DNA-binding domain. Instead, the mutation might affect the ability of SIX6 to interact with other transcription factors or with co-factor proteins, altering the efficiency with which the protein can activate its target genes.

To identify possible target genes for SIX6, Skowronska-Krawczyk and colleagues again turned to genetic-association studies. These indicated that mutations in the *p16INK4a* gene are a strong risk factor for glaucoma. The authors found that expression of both *p16INK4a* and *SIX6* was higher in eyes of people with glaucoma than in those of healthy people. Moreover, they demonstrated that SIX6 binds to and activates *p16INK4a*.

In many cell types, *p16INK4a* is associated with a cellular ageing process called senescence. Skowronska-Krawczyk *et al.* found that approximately four times more RGCs were senescing in patients with glaucoma than in healthy people. To probe this pathway further, the authors engineered human retinal progenitor cells cultured *in vitro* to express the SIX6 His141 mutation. The mutant protein strongly upregulated *p16INK4a* and another marker of cellular senescence, the *IL-6* gene. This effect seems to be specific to the His141 mutation, because upregulation of these markers did not occur in cells producing wild-type SIX6 or forms of SIX6 mutated at different residues. Together, the results indicate that the His141 mutation increases the effectiveness with which SIX6 activates *p16INK4a* and triggers senescence pathways in RGCs.

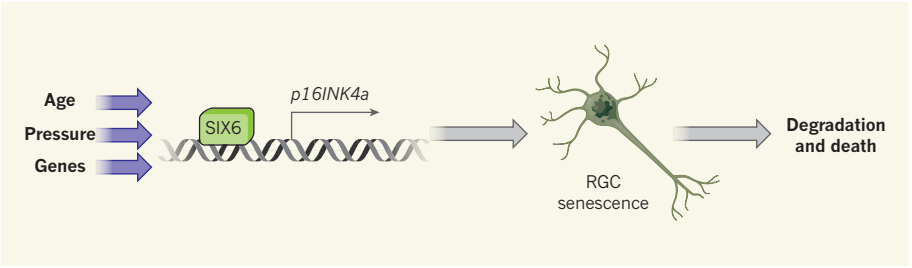


Figure 1 | Molecular pathways that underlie glaucoma. Age, elevated pressure in the eye and certain genetic mutations are all associated with an increased risk of glaucoma, a form of blindness linked to the degradation of retinal ganglion cells (RGCs). Skowronska-Krawczyk *et al.*[4] report that these risk factors converge on a single molecular cascade in which the transcription factor SIX6 binds to and activates the gene *p16INK4a*. Increased *p16INK4a* expression causes RGC senescence and, eventually, RGC degradation and death.

Skowronska-Krawczyk and colleagues next explored whether activation of *p16INK4a* was linked to RGC ageing or death in mice in which intraocular pressure was experimentally raised. They found that expression of both *SIX6* and *p16INK4a* increased markedly after experimental elevation of intraocular pressure. The evidence for an interaction between SIX6 and *p16INK4a* was further bolstered by the discovery that *p16INK4a* expression was reduced in mice lacking SIX6, and that elevated intraocular pressure increased SIX6–*p16INK4a* binding in wild-type mice. As in human glaucomatous retinas, increases in intraocular pressure dramatically elevated the number of senescent RGCs. Together, these results suggest that increased *p16INK4a* expression is a major cause of cellular-senescence pathways that ultimately lead to RGC degeneration and death in glaucoma.

In a final set of experiments, the authors performed a crucial test of this model by assessing whether genetic deletion of *p16INK4a* or partial deletion of *SIX6* impeded RGC death in a mouse model of glaucoma. Remarkably, when intraocular pressure was experimentally increased in either of these genetically mutated mouse strains, RGCs resisted death, strongly supporting the idea that SIX6-activated increases in *p16INK4a* mediate RGC loss in response to different stressors (Fig. 1).

Skowronska-Krawczyk and colleagues' study is an important step forward. First, it provides support for the long-held view that, even though different risk factors and stressors can increase the likelihood of glaucoma, there is a common molecular mechanism by which those stressors act to kill RGCs. Second, the study indicates that cellular senescence and its associated pathways are precursors to RGC degeneration and death.

Over the past few years, there has been a surge in our understanding about which RGCs are most vulnerable in early-stage glaucoma[8,9], and of the ion channels required to translate intraocular pressure increases into RGC degradation and death[10]. The current study provides a solid molecular foundation on which to integrate these findings. A more complete understanding of the biological underpinnings of glaucoma will no doubt also help to identify new targets for intervention, and might reveal mechanistic insights into the molecular basis of other age-related neurodegenerative diseases, such as Alzheimer's and Parkinson's disease. ■

Andrew D. Huberman *and* **Rana N. El-Danaf** *are in the Neurobiology Section, Division of Biological Science, and in the Departments of Neurosciences and Ophthalmology, School of Medicine, University of California, San Diego, La Jolla, California 92093, USA.* **A.D.H.** *is also at the Salk Institute for Biological Studies, La Jolla.*
e-mails: ahuberman@ucsd.edu; reldanaf@ucsd.edu

1. Dhande, O. S. & Huberman, A. D. *Curr. Opin. Neurobiol.* **24,** 133–142 (2014).
2. Kwon, Y. H., Fingert, J. H., Kuehn, M. H. & Alward, W. L. M. *N. Engl. J. Med.* **360,** 1113–1124 (2009).
3. Weinreb, R. N., Aung, T. & Medeiros, F. A. *J. Am. Med. Assoc.* **311,** 1901–1911 (2014).
4. Skowronska-Krawczyk, D. *et al. Mol. Cell* **59,** 931–940 (2015).
5. Jain, S. & Aref, A. A. *J. Ophthalmic Vis. Res.* **10,** 178–183 (2015).
6. Tielsch, J. M. *et al. J. Am. Med. Assoc.* **266,** 369–374 (1991).
7. Anderson, A. M., Weasner, B. M., Weasner, B. P. & Kumar, J. P. *Development* **139,** 991–1000 (2012).
8. Della Santina, L., Inman, D. M., Lupien, C. B., Horner, P. J. & Wong, R. O. *J. Neurosci.* **33,** 17444–17457 (2013).
9. El-Danaf, R. N. & Huberman, A. D. *J. Neurosci.* **35,** 2329–2343 (2015).
10. Ward, N. J., Ho, K. W., Lambert, W. S., Weitlauf, C. & Calkins, D. J. *J. Neurosci.* **34,** 3161–3170 (2014).

CORRECTION

The News & Views article 'Rehabilitation: Boost for movement' by Randolph J. Nudo (*Nature* **527,** 314–315; 2015) omitted to mention that the author has declared competing financial interests. Details are available in the online version of the article.

ARTICLE

OPEN

doi:10.1038/nature16150

Hemichordate genomes and deuterostome origins

Oleg Simakov[1,2]*, Takeshi Kawashima[3]†*, Ferdinand Marlétaz[4], Jerry Jenkins[5], Ryo Koyanagi[6], Therese Mitros[7], Kanako Hisata[3], Jessen Bredeson[7], Eiichi Shoguchi[3], Fuki Gyoja[3], Jia-Xing Yue[8]†, Yi-Chih Chen[9], Robert M. Freeman Jr[10]†, Akane Sasaki[11], Tomoe Hikosaka-Katayama[12], Atsuko Sato[13], Manabu Fujie[6], Kenneth W. Baughman[3], Judith Levine[14], Paul Gonzalez[14], Christopher Cameron[15], Jens H. Fritzenwanker[14], Ariel M. Pani[16], Hiroki Goto[6], Miyuki Kanda[6], Nana Arakaki[6], Shinichi Yamasaki[6], Jiaxin Qu[17], Andrew Cree[17], Yan Ding[17], Huyen H. Dinh[17], Shannon Dugan[17], Michael Holder[17], Shalini N. Jhangiani[17], Christie L. Kovar[17], Sandra L. Lee[17], Lora R. Lewis[17], Donna Morton[17], Lynne V. Nazareth[17], Geoffrey Okwuonu[17], Jireh Santibanez[17], Rui Chen[17], Stephen Richards[17], Donna M. Muzny[17], Andrew Gillis[18], Leonid Peshkin[10], Michael Wu[7], Tom Humphreys[19], Yi-Hsien Su[9], Nicholas H. Putnam[8]†, Jeremy Schmutz[5], Asao Fujiyama[20], Jr-Kai Yu[9], Kunifumi Tagawa[11], Kim C. Worley[17], Richard A. Gibbs[17], Marc W. Kirschner[10], Christopher J. Lowe[14], Noriyuki Satoh[3], Daniel S. Rokhsar[1,7,21] & John Gerhart[7]

Acorn worms, also known as enteropneust (literally, 'gut-breathing') hemichordates, are marine invertebrates that share features with echinoderms and chordates. Together, these three phyla comprise the deuterostomes. Here we report the draft genome sequences of two acorn worms, *Saccoglossus kowalevskii* and *Ptychodera flava*. By comparing them with diverse bilaterian genomes, we identify shared traits that were probably inherited from the last common deuterostome ancestor, and then explore evolutionary trajectories leading from this ancestor to hemichordates, echinoderms and chordates. The hemichordate genomes exhibit extensive conserved synteny with amphioxus and other bilaterians, and deeply conserved non-coding sequences that are candidates for conserved gene-regulatory elements. Notably, hemichordates possess a deuterostome-specific genomic cluster of four ordered transcription factor genes, the expression of which is associated with the development of pharyngeal 'gill' slits, the foremost morphological innovation of early deuterostomes, and is probably central to their filter-feeding lifestyle. Comparative analysis reveals numerous deuterostome-specific gene novelties, including genes found in deuterostomes and marine microbes, but not other animals. The putative functions of these genes can be linked to physiological, metabolic and developmental specializations of the filter-feeding ancestor.

The prominent pharyngeal gill slits, rigid stomochord, and midline nerve cords of acorn worms led 19th century zoologists to designate them as 'hemichordates' and group them with vertebrates and other chordates[1–4], but their early embryos and larvae also linked them to echinoderms[5,6]. Current molecular phylogenies strongly support the affinities of hemichordates and echinoderms as sister phyla, together called ambulacrarians[7], and unite ambulacrarians and chordates within the deuterostomes (see glossary in Supplementary Note 1). Of all the shared derived morphological characters proposed between hemichordates and chordates, the pharyngeal gill slits have emerged with unambiguous morphological and molecular support, notably the shared expression of the *pax1/9* gene[8–10]. These structures were ancestral deuterostome characters elaborated upon the bilaterian ancestral body plan, but the gill slits were subsequently lost in extant echinoderms and amniotes[11]. Since extant invertebrate deuterostomes use this apparatus for efficient suspension and/or deposit feeding, the early Cambrian or Precambrian deuterostome ancestor probably also shared this lifestyle. This perspective on the last common deuterostome ancestor informs our understanding of the subsequent evolution of hemichordates, echinoderms and chordates[10,12–16].

Hemichordates share bilateral symmetry, gill slits, soft bodies and early axial patterning with chordates, making them key comparators for inferring the ancestral genomic features of deuterostomes. To this end, we sequenced and analysed the genomes of acorn worms belonging to the two main lineages of enteropneust hemichordates (Supplementary Note 1): *Saccoglossus kowalevskii* (Harrimaniidae; Atlantic, North America, Fig. 1a) and *Ptychodera flava* (Ptychoderidae; Pacific, pan-tropical, Fig. 1b). Both have characteristic three-part

[1]Molecular Genetics Unit, Okinawa Institute of Science and Technology Graduate University, Onna, Okinawa 904-0495, Japan. [2]Department of Molecular Evolution, Centre for Organismal Studies, University of Heidelberg, 69115 Heidelberg, Germany. [3]Marine Genomics Unit, Okinawa Institute of Science and Technology Graduate University, Onna, Okinawa 904-0495, Japan. [4]Department of Zoology, University of Oxford, Oxford OX1 3PS, UK. [5]HudsonAlpha Institute of Biotechnology, Huntsville, Alabama 35806, USA. [6]DNA Sequencing Section, Okinawa Institute of Science and Technology Graduate University, Onna, Okinawa 904-0495, Japan. [7]Department of Molecular and Cell Biology, University of California, Berkeley California 94720-3200, USA. [8]Department of Ecology and Evolutionary Biology, Rice University, Houston, Texas 77005, USA. [9]Institute of Cellular and Organismic Biology, Academia Sinica, Taipei 11529, Taiwan. [10]Department of Systems Biology, Harvard Medical School, Boston, Massachusetts 02115, USA. [11]Marine Biological Laboratory, Graduate School of Science, Hiroshima University, Onomichi, Hiroshima 722-0073, Japan. [12]Natural Science Center for Basic Research and Development, Gene Science Division, Hiroshima University, Higashi-Hiroshima, Hiroshima 739-8527, Japan. [13]Marine Biological Association of the UK, The Laboratory, Citadel Hill, Plymouth PL1 2PB, UK. [14]Department of Biology, Hopkins Marine Station, Stanford University, Pacific Grove, California 93950, USA. [15]Département de sciences biologiques, University of Montreal, Quebec H3C 3J7, Canada. [16]University of North Caroline at Chapel Hill, North Carolina 27599, USA. [17]Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, MS BCM226, Houston, Texas 77030, USA. [18]Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK. [19]Institute for Biogenesis Research, University of Hawaii, Hawaii 96822, USA. [20]National Institute of Genetics, Mishima, Shizuoka 411-8540, Japan. [21]US Department of Energy Joint Genome Institute, Walnut Creek, California 94598, USA. †Present addresses: University of Tsukuba, Tsukuba, Ibaraki 305-572, Japan (T.K.); Institute for Research on Cancer and Aging, Nice (IRCAN), CNRS UMR 7284, INSERM U 1081, Nice 06107, France (J.-X.Y.); FAS Research Computing, Harvard University, Cambridge, Massachusetts 02138, USA (R.M.F.); Dovetail Genomics, Santa Cruz, California 95060, USA (N.H.P.).

*These authors contributed equally to this work.

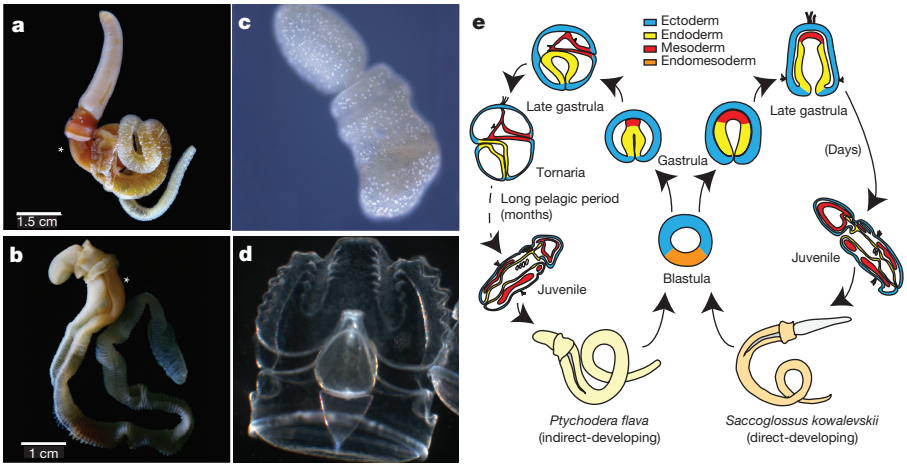


Figure 1 | Hemichordate model systems and their embryonic development. The hemichordate phylum includes the enteropneusts (acorn worms) and pterobranchs (minute, colonial, tube-dwelling; not shown). **a**, **c**, *Saccoglossus kowalevskii* (Harrimaniid (direct developing) enteropneust) adult (**a**) and juvenile (**c**) with gill slits. **b**, **d**, *Ptychodera flava* (Ptychoderid (indirect developing) enteropneust) adult (**b**) and the tornaria stage larva (**d**). Gill slits labelled with an asterisk in **a** and **b**. **e**, Comparison of the direct and indirect modes of development of the two hemichordates, indicating the long pelagic larval period in *Ptychodera* until the settlement and metamorphosis as a juvenile.

bodies comprising proboscis, collar and trunk, the last with tens to hundreds of pairs of gill slits. While *S. kowalevskii* develops directly to a juvenile worm with these traits within days (Fig. 1c, e), *P. flava* develops indirectly through a feeding larva that metamorphoses to a juvenile worm after months in the plankton (Fig. 1d, e). Our analyses begin to integrate macroscopic information about morphology, organismal physiology, and descriptive embryology of these deuterostomes with genomic information about gene homologies, gene arrangements, gene novelties and non-coding elements.

Genomes

We sequenced the two acorn worm genomes by random shotgun methods with a variety of read types (Methods; Supplementary Note 2), each starting from sperm from a single outbred diploid individual. The haploid lengths of the two genomes are both about 1 Gbp (Extended Data Fig. 1), but differ in nucleotide heterozygosity. Both acorn worm genomes were annotated using extensive transcriptome data as well as standard homology-based and *de novo* methods (Supplementary Note 3). Counting gene models with at least one detectable orthologue in another sequenced metazoan species, we find that *Ptychodera* and *Saccoglossus* encode at least 18,556 and 19,270 genes, respectively (Methods). Additional *de novo* gene predictions include divergent and/or novel genes (Extended Data Fig. 1). Despite the ancient divergence of the *Saccoglossus* and *Ptychodera* lineages (more than 370 million years ago, see below) and their different modes of development, the two acorn worm genomes have similar bulk gene content, as discussed later (Extended Data Fig. 2 and Supplementary Note 4), and similar repetitive landscapes (Supplementary Note 5).

Deuterostome phylogeny

Deuterostome relationships were originally inferred from developmental and morphological characters[2,5,17] and these hypotheses were later tested and refined with molecular data[6,7]. Aspects of deuterostome phylogeny continue to be controversial, however, notably the position of the sessile pterobranchs among hemichordates, and the surprising association of *Xenoturbella*[18] and acoelomorph flatworms with ambulacrarians[19] proposed by some studies. We explored these issues using genome-wide analyses of the newly sequenced hemichordate genomes augmented with extensive new RNAseq from five echinoderms, three additional hemichordates (including a rhabdopleurid pterobranch) and two acoels (Fig. 2, Extended Data Fig. 3, Methods and Supplementary Note 6). We recovered the monophyly of hemichordates, echinoderms, ambulacrarians and deuterostomes, using not only amino acid characters but also presence–absence characters for introns and coding indels (Supplementary Note 4). Our analyses also placed pterobranch hemichordates as the sister-group to enteropneusts[7] rather than within them[12]. These phylogenetic analyses imply that genomic traits shared by chordates and ambulacrarians can be attributed to the last common deuterostome ancestor (see below). Using a relaxed molecular clock, we estimate a Cambrian origin of hemichordates (Methods, Extended Data Fig. 3 and Supplementary Note 6).

We also performed several analyses to assess the controversial relationships between *Xenoturbella*, acoelomorphs and deuterostomes (Supplementary Note 6). With conventional site-homogeneous models, acoels remain outside deuterostomes[20–23] (Fig. 2, Supplementary Figs 6.1 and 6.2). Alternative models[24], however, show equivocal branching of acoels depending on the inclusion of the current sparse data for *Xenoturbella* (Supplementary Note 6). Notably, without *Xenoturbella*, acoels are positioned as a bilaterian sister group (Supplementary Fig. 6.3)[20–23]. Although we cannot rule out a deuterostome placement for *Xenoturbella*, our analyses generally do not support a grouping of acoels with deuterostomes[19].

The gene set of the deuterostome ancestor

By comparative analysis, we identified 8,716 families of homologous genes whose distributions in sequenced extant genomes imply their presence in the deuterostome ancestor (Methods; Supplementary Note 4). Owing to gene duplication and other processes the descendants of these ancestral genes account for ~14,000 genes in extant deuterostome genomes including human (Supplementary Table 4.1.2). The distributions of gene functions, domain compositions, and gene family sizes of hemichordates resemble those of amphioxus, sea urchin, and sequenced lophotrochozoans more than those of ecdysozoans; vertebrates also form a distinct group (Extended Data Fig. 2, Supplementary Note 4 and Supplementary Fig. 4.2).

Exon–intron structures of genes are generally well conserved among hemichordates, chordates, and many non-deuterostome metazoans, allowing us to infer 2,061 ancestral deuterostome splice sites (Supplementary Note 4). Among orthologous bilaterian genes we found 23 introns and 4 coding sequence indels present only in deuterostomes (shared between at least one ambulacrarian and chordate), suggesting that these shared derived characters may be useful to diagnose clade membership of new candidate organisms (Supplementary Note 4).

Based on whole-genome alignments, we identified 6,533 conserved non-coding elements (CNE) longer than 50 bp that are found in all of the five deuterostomes *Saccoglossus*, *Ptychodera*, amphioxus,

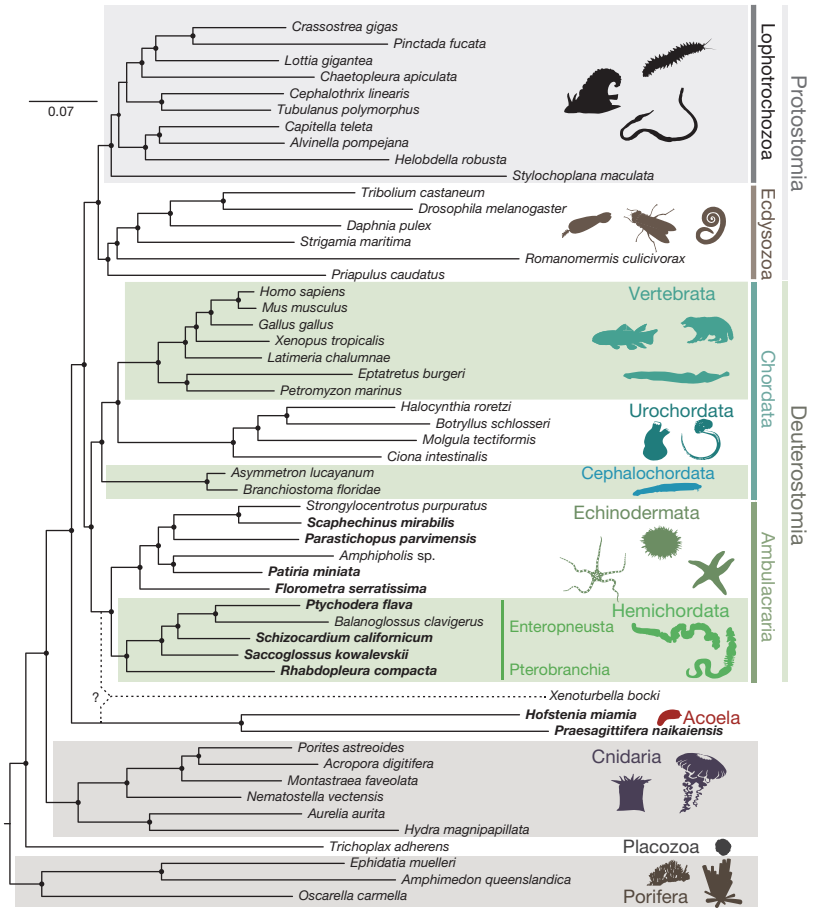


Figure 2 | Phylogenetic placement of deuterostome taxa within the metazoan tree. Maximum-likelihood tree obtained with a super-matrix of 506,428 amino-acid residues gathered from 1,564 orthologous genes in 52 species (65.1% occupancy) and using a LG+Γ model partitioned for each gene. Filled circles at nodes denote maximal bootstrap support. Taxa highlighted in bold are newly sequenced genomes and transcriptomes introduced in this study. Bar indicates the number of substitutions per site.

sea urchin, and human (Methods; Supplementary Note 8). The identified CNEs overlap extensively with human long non-coding RNAs (3,611 CNE loci; 55%, Fisher's exact test P value $< 2.2 \times 10^{-16}$). Those alignments usually do not exceed 250 bp (as has been reported among vertebrates[25]) and occur in clusters (Supplementary Note 8). Among these conserved sequences is a previously identified vertebrate brain and neural tube specific enhancer, located close to the *sox14/21* orthologue in all five species[26].

Conserved gene linkage

Ancient gene linkages ('macro-synteny'[27]) are often preserved in extant bilaterian genomes[27,28]. Comparative analysis revealed 17 ancestral linkage groups across chordates, including amphioxus and *Ciona*[27]. While the contiguity of the draft of the sea urchin genome assembly[29] is too limited to determine whether it shares this chromosome-scale organization, we find that the *Saccoglossus* genome clearly shares these chordate-defined linkage groups (Fig. 3a and Supplementary Note 7), implying that these chromosome-scale linkages were also present in the ancestral deuterostome.

On a more local scale, we find hundreds of tightly linked conserved gene clusters of three or more genes ('micro-synteny'; Methods; Supplementary Note 7) including Hox[30] and ParaHox[31] clusters in both acorn worms (Extended Data Fig. 4), as also found in echinoderms[32,33]. *Saccoglossus* and amphioxus share more micro-syntenic linkages with each other than either does with sea urchin, vertebrates, or available protostome genomes (Methods, Fig. 3b and Extended Data Figs 5 and 6). Conservation of micro-syntenic linkages can occur due to low rates of genomic rearrangement or, more interestingly, as a result of selection to retain linkages between genes and their regulatory elements located in neighbouring genes[28].

A deuterostome pharyngeal gene cluster

One conserved deuterostome-specific micro-syntenic cluster with functional implications for deuterostome biology is a cluster of genes expressed in the pharyngeal slits and surrounding pharyngeal endoderm (Fig. 4; Supplementary Note 9). This six-gene cluster contains four transcription factor genes in the order *nkx2.1*, *nkx2.2*, *pax1/9* and *foxA*, along with two non-transcription-factor genes *slc25A21* and *mipol1*, whose introns harbour regulatory elements for *pax1/9* and *foxA*, respectively[34–36]. The cluster was first found conserved across vertebrates including humans (see chromosome 14; 1.1 Mb length from *nkx2.1* to *foxA1*)[34,37]. In *S. kowalevskii*, it is intact with the same gene order as in vertebrates (0.5 Mb length from *nkx2.1* to *foxA*), implying that it was present in the deuterostome and ambulacrarian ancestors. The full ordered gene cluster also exists on a single scaffold in the crown-of-thorns sea star *Acanthaster planci*. Since these genes are not clustered in available protostome genomes, there is no evidence for deeper bilaterian ancestry. Two non-coding elements that are conserved across vertebrates and amphioxus[38] are found in the hemichordate and *A. planci* clusters at similar locations (A2 and A4, in Fig. 4a).

The *pax1/9* gene, at the centre of the cluster, is expressed in the pharyngeal endodermal primordium of the gill slit in hemichordates, tunicates, amphioxus, fish, and amphibians[8,9], and in the branchial pouch endoderm of amniotes (which do not complete the last steps of gill slit formation), as well as other locations in vertebrates. The *nkx2.1* (thyroid transcription factor 1) gene is also expressed in the hemichordate

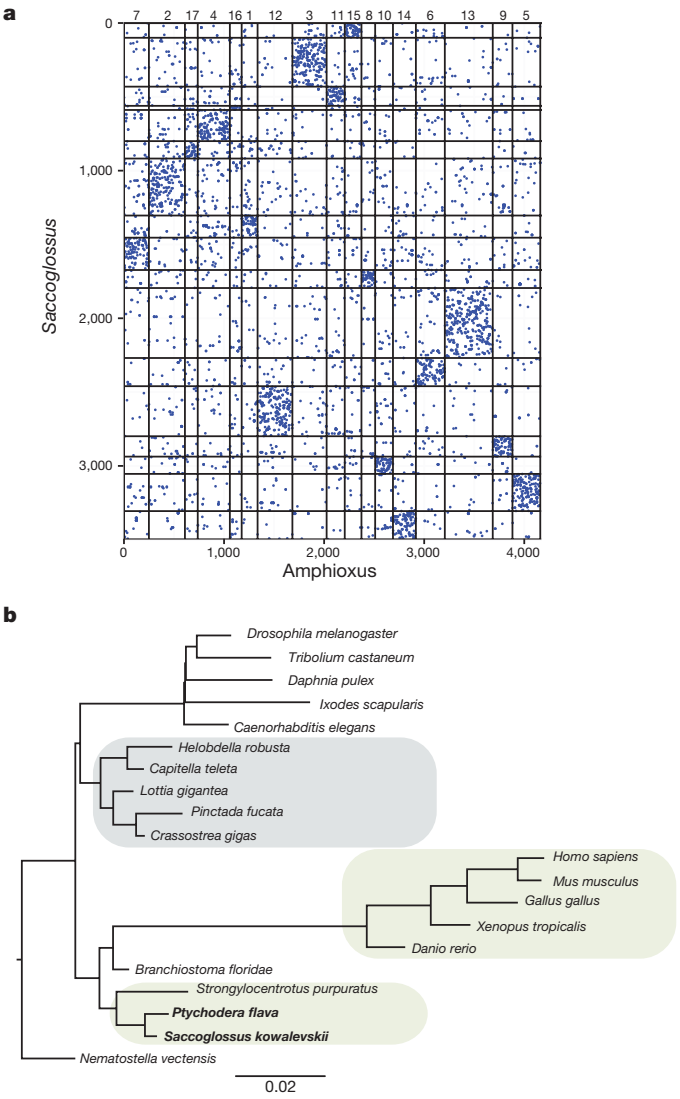


Figure 3 | High level of linkage conservation in *Saccoglossus*. **a**, Macro-synteny dot plot between *Saccoglossus* and amphioxus; each dot represents two orthologous genes linked in the two species, and ordered according to their macro-syntenic linkage. Amphioxus scaffolds are organized according to the 17 ancestral linkage groups (ALGs) inferred by comparison of the amphioxus and vertebrate genomes[27]. Intersection areas of highest dot density are marked by numbers along the top of the plot, identifying each of the 17 putative ALGs. Axes represent orthologous gene group index along the genome. **b**, Branch-length estimation for loss and gain of synteny blocks with MrBayes, see Supplementary Note 7 for details. Short branches in hemichordates (in bold) indicate a high level of micro-syntenic retention in their genomes.

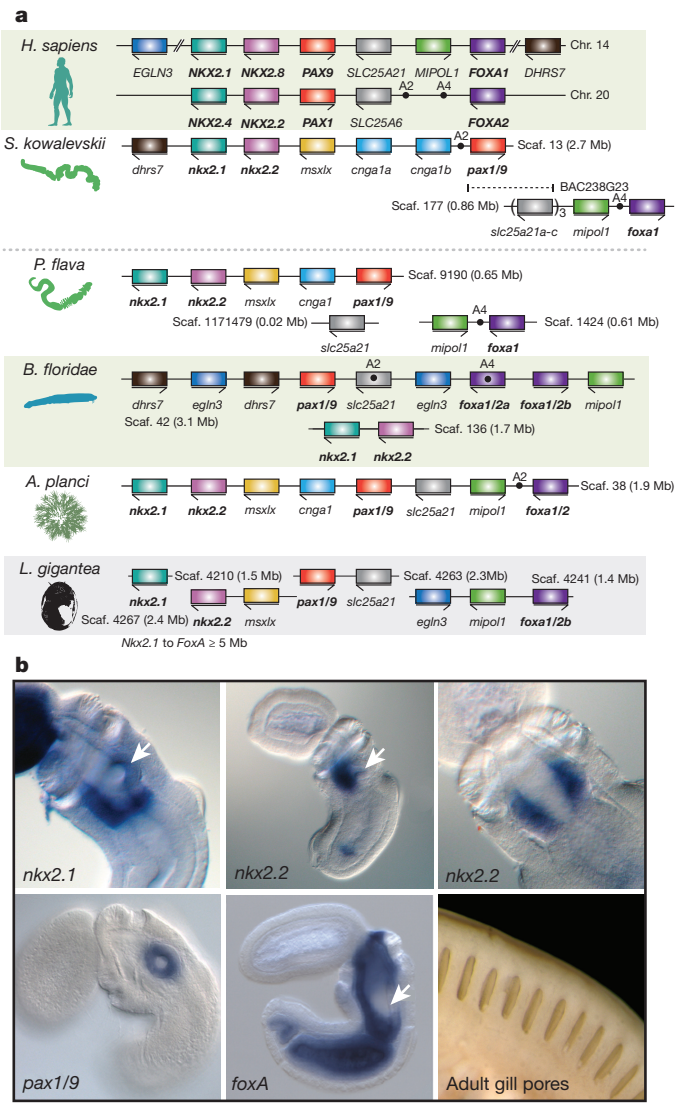


Figure 4 | Conservation of a pharyngeal gene cluster across deuterostomes. **a**, Linkage and order of six genes including the four genes encoding transcription factors Nkx2.1, Nkx2.2, Pax1/9 and FoxA, and two genes encoding non-transcription factors Slc25A21 (solute transporter) and Mipol1 (mirror-image polydactyly 1 protein), which are putative 'bystander' genes containing regulatory elements of *pax1/9* and *foxA*, respectively. The pairings of *slc25A21* with *pax1/9* and of *mipol1* with *foxA* occur also in protostomes, indicating bilaterian ancestry. The cluster is not present in protostomes such as *Lottia* (Lophotrochozoa), *Drosophila melanogaster*, *Caenorhabditis elegans* (Ecdysozoa), or in the cnidarian, *Nematostella*. *SLC25A6* (the *slc25A21* paralogue on human chromosome 20) is a potential pseudogene. The dots marking A2 and A4 indicate two conserved non-coding sequences first recognized in vertebrates and amphioxus[36], also present in *S. kowalevskii* and, partially, in *P. flava* and *A. planci*. **b**, The four transcription factor genes of the cluster are expressed in the pharyngeal/foregut endoderm of the *Saccoglossus* juvenile: *nkx2.1* is expressed in a band of endoderm at the level of the forming gill pore, especially ventral and posterior to it (arrow), and in a separate ectodermal domain in the proboscis. It is also known as thyroid transcription factor 1 due to its expression in the pharyngeal thyroid rudiment in vertebrates. The *nkx2.2* gene is expressed in pharyngeal endoderm just ventral to the forming gill pore, shown in side view (arrow indicates gill pore) and ventral view; and *pax1/9* is expressed in the gill pore rudiment itself. In *S. kowalevskii*, this is its only expression domain, whereas in vertebrates it is also expressed in axial mesoderm. The *foxA* gene is expressed widely in endoderm but is repressed at the site of gill pore formation (arrow). An external view of gill pores is shown; up to 100 bilateral pairs are present in adults, indicative of the large size of the pharynx.

pharyngeal endoderm in a band passing through the gill slit, but not localized to a thyroid-like organ[39]. Here we also examined the expression of *nkx2.2* and *foxA* in *S. kowalevskii*. We find that *nkx2.2*, which is expressed in the ventral hindbrain in vertebrates, is expressed in pharyngeal ventral endoderm in *S. kowalevskii*, close to the gill slit (Fig. 4b), and that *foxA* is expressed throughout endoderm but repressed in the gill slit region (Fig. 4b). The co-expression of this ordered cluster of the four transcription factors during pharyngeal development strongly supports the functional importance of their genomic clustering.

The presence of this cluster in the crown-of-thorns sea star, an echinoderm that lacks gill pores, and in amniote vertebrates that lack gill slits, suggests that the cluster's ancestral role was in pharyngeal apparatus patterning as a whole, of which overt slits (perforations of

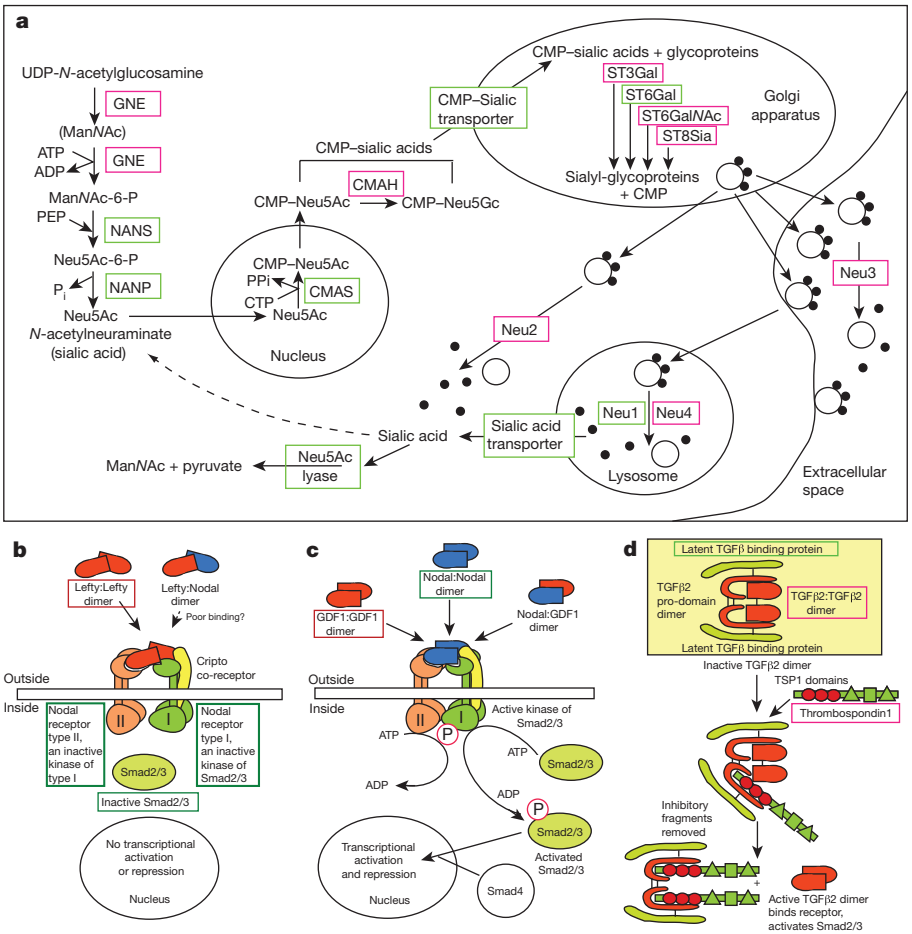


Figure 5 | Examples of deuterostome gene novelties. a, Steps of biosynthesis of sialic acid and its addition to and removal from glycoproteins. **b**–**d**, Novel genes in TGFβ signalling pathways. The encoded proteins are shown and include Lefty (**b**), an antagonist of Nodal signalling, which activates Smad2/3-dependent transcription when not antagonized; Univin (**c**), an agonist of Nodal signalling, also called Vg1, DVR1, and GDF1; and TGFβ2 (**d**), a ligand that activates Smad2/3-dependent transcription by binding to a deuterostome-specific TGFβ receptor type II, which contains a novel ectodomain (not shown). Also shown in **d** is the novel protein thrombospondin 1 that activates TGFβ2 by releasing it from an inactive complex, by way of its TSP1 domains. Red boxes around protein names indicate their deuterostome novelty. Green boxes around the names indicate genes with pan-metazoan/bilaterian ancestry and without accelerated sequence change in the deuterostome lineage.

apposed endoderm and ectoderm) were but one part, and the cluster is retained in these cases because of its continuing contribution to pharynx development. Genomic regions of the pharyngeal cluster have been implicated in long-range promoter–enhancer interactions, supporting the regulatory importance of this gene linkage (see Supplementary Note 9)[40]. Alternatively, genome rearrangement in these lineages may be too slow to disrupt the cluster even without functional constraint. Here we propose that the clustering of the four ordered transcription factors, and their bystander genes, on the deuterostome stem served a regulatory role in the evolution of the pharyngeal apparatus, the foremost morphological innovation of deuterostomes.

Deuterostome novelties

We found >30 deuterostome genes with sequences that differ markedly from those of other metazoans, related to functional innovation in deuterostomes. Some plausibly arose from accelerated sequence change on the deuterostome stem from distant but identifiable bilaterian homologues, others represent new protein domain combinations in deuterostomes, while others lack identifiable sequence and domain homologues in other animals. In the latter group, we found over a dozen deuterostome genes that have readily identified relatives in marine microbes, often cyanobacteria or eukaryotic micro-algae, but are not known in other metazoans (Extended Data Table 1 and Extended Data Fig. 7; Supplementary Notes 10.4 and 10.5). Such genes include two of the novel deuterostome sequences associated with sialic acid metabolism (found in many microbes[41], see below), enzymes that modify proteins (for example, protein arginine deiminase) and RNA (for example, FATSO methyladenosine demethylase) as well as others that provide specialized reactions of secondary metabolism (Extended Data Table 1 and Extended Data Fig. 7; Supplementary Note 10.5). Possible explanations for the unusual phylogenetic distribution of these genes include horizontal transfer on the deuterostome stem from early marine microbes (which were plausibly commensals, pathogens, or food sources of stem deuterostomes), or convergent gene loss and/or extensive sequence divergence along five or more opisthokont lineages (Supplementary Note 10.2).

Regardless of their mechanism of origination, the various deuterostome novelties and gene family expansions of sialic acid metabolism are noteworthy. Deuterostomes are unique among metazoans in their high level and diverse linkage of addition of sialic acid (also known as neuraminic acid), a nine carbon negatively charged sugar, to the terminal sugars of glycoproteins, mucins and glycolipids[42]. We find expanded families of enzymes for several of these reactions in hemichordates (Fig. 5a and Extended Data Table 1). Based on the presence/absence of relevant enzymes we infer that 5 of the 11 steps of the pathways of sialic acid formation, addition to termini, and removal are not found in protostomes or other metazoans, and are deuterostome novelties (Fig. 5a and Supplementary Note 10), whereas the other steps use

enzymes similar to those of the more limited pathway of some protostomes (for example, insects such as *Drosophila*)[43].

The importance of glycoproteins for muco-ciliary feeding and other hemichordate activities is further supported by novel and expanded families of genes encoding the polypeptide backbones of glycoproteins, those with von Willebrand type-D and/or cysteine-rich domains (PTHR11339 classifier), including mucins, present in hemichordates and amphioxus as large tandemly duplicated clusters (with varied expression patterns as shown in Extended Data Fig. 8), but not in sea urchin, which has a different mode of feeding (Supplementary Note 10). As in amphioxus, the pharynx of *Saccoglossus* is heavily ciliated[44,45], and cells of the pharyngeal walls in hemichordates and the ventral endostyle in amphioxus secrete abundant mucins and glycoproteins[46]. Similarly, in the deuterostome ancestor these glycoproteins probably enhanced the muco-ciliary filter-feeding capture of food particles from the microbe-rich marine environment and protected its inner and outer tissue surfaces.

Novelty in the TGFβ signalling pathway

The signalling ligands Lefty (a Nodal antagonist) and Univin/Vg1/GDF1[47] (a Nodal agonist) are deuterostome innovations that modulate Nodal signalling during the major developmental events of endomesoderm induction and axial patterning in vertebrates, axial patterning in hemichordates and echinoderms, and left–right patterning in all deuterostomes[48] (see Fig. 5b–d and Extended Data Fig. 9a, b). *Univin* is tightly linked to the related bilaterian *bmp2/4* in the sea urchin genome[49] and also, we now report, in hemichordates and amphioxus, supporting its origin by tandem duplication and divergence from an ancestral *bmp2/4*-type gene, as suggested previously[49].

TGFβ2 signalling (TGFβ1, 2 and 3 in vertebrates) is a deuterostome innovation that controls cell growth, proliferation, differentiation and apoptosis at later developmental stages. Accompanying the novel TGFβ2 ligand, the type II receptor has a novel ectodomain. The extracellular matrix protein thrombospondin 1, which activates TGFβ2 in vertebrates, contains a deuterostome-unique combination of domains including three thrombospondin type 1 (TSP1) domains that bind the TGFβ2 pro-domain region. While these signalling novelties have clear sequence similarity to pan-bilaterian components, they form long stem branch clades on the phylogenetic trees, indicating extensive sequence divergence on the deuterostome stem (Supplementary Note 10). Together, these innovations appear to contribute to the increased amount and complex patterning of Smad2/3-mediated signalling in deuterostomes compared with protostomes and other metazoans.

Conclusion

The two acorn worms whose genomes are described here represent the two main enteropneust lineages, separated by at least 370 million years and differing in their developmental modes. These analyses reveal (1) extensive conserved macro-synteny among deuterostomes; (2) a widely conserved deuterostome-specific cluster of six ordered genes, including four transcription factor genes that are expressed during the development of pharyngeal gill slits and the branchial apparatus, the most prominent morphological innovation of the deuterostome ancestor; and (3) numerous gene novelties shared among deuterostomes, many expanded into large families, with putative protein functions that imply physiological, metabolic and developmental specializations of the filter-feeding deuterostome ancestor. Some of these genes lack identifiable orthologues in other metazoans but do resemble microbial sequences and domain types. In addition to their contributions towards defining the deuterostome ancestor and illuminating chordate origins, the two genomes should inform hypotheses of larval evolution by providing a basis for future comparisons of direct-developing and indirect-developing acorn worms, which achieve remarkably similar adult forms by distinct embryological routes (Fig. 1).

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 15 July; accepted 13 October 2015.
Published online 18 November 2015.

1. Bateson, W. The later stages in the development of *Balanoglossus Kowalevskii*, with a suggestion as to the affinities of the Enteropneusta. 2 parts. *Q. J. Microsc. Sci.* **25,** 81–122 (1885).
2. Bateson, W. Memoirs: the ancestry of the Chordata. *Q. J. Microsc. Sci.* **2,** 535–572 (1886).
3. Kovalevskij, A. O. *Anatomie des Balanoglossus delle Chiaje* (*Mémoires de l'Académie Impériale des Sciences de St. Pétersbourg: Imperatorskaja Akademija Nauk*, 1866).
4. Agassiz, A. The history of *Balanoglossus* and tornaria. *Memoirs of the American Academy of Arts and Sciences* **9,** 421–436 (1873).
5. Metschnikoff, V. Über die systematische Stellung von Balanoglossus. *Zool. Anz.* **4,** 139–157 (1881).
6. Halanych, K. M. The phylogenetic position of the pterobranch hemichordates based on 18S rDNA sequence data. *Mol. Phylogenet. Evol.* **4,** 72–76 (1995).
7. Cannon, J. T. *et al.* Phylogenomic resolution of the hemichordate and echinoderm clade. *Curr. Biol.* **24,** 2827–2832 (2014).
8. Ogasawara, M., Wada, H., Peters, H. & Satoh, N. Developmental expression of *Pax1/9* genes in urochordate and hemichordate gills: insight into function and evolution of the pharyngeal epithelium. *Development* **126,** 2539–2550 (1999).
9. Gillis, J. A., Fritzenwanker, J. H. & Lowe, C. J. A stem-deuterostome origin of the vertebrate pharyngeal transcriptional network. *Proc. R. Soc. Lond. B* **279,** 237–246 (2012).
10. Lowe, C. J., Clarke, D. N., Medeiros, D. M., Rokhsar, D. S. & Gerhart, J. The deuterostome context of chordate origins. *Nature* **520,** 456–465 (2015).
11. Swalla, B. J. & Smith, A. B. Deciphering deuterostome phylogeny: molecular, morphological and palaeontological perspectives. *Phil. Trans. R. Soc. Lond. B* **363,** 1557–1568 (2008).
12. Cameron, C. B., Garey, J. R. & Swalla, B. J. Evolution of the chordate body plan: new insights from phylogenetic analyses of deuterostome phyla. *Proc. Natl Acad. Sci. USA* **97,** 4469–4474 (2000).
13. Gerhart, J., Lowe, C. & Kirschner, M. Hemichordates and the origin of chordates. *Curr. Opin. Genet. Dev.* **15,** 461–467 (2005).
14. Gonzalez, P. & Cameron, C. B. The gill slits and pre-oral ciliary organ of *Protoglossus* (Hemichordata: Enteropneusta) are filter-feeding structures. *Biol. J. Linn. Soc.* **98,** 898–906 (2009).
15. Brown, F. D., Prendergast, A. & Swalla, B. J. Man is but a worm: chordate origins. *Genesis* **46,** 605–613 (2008).
16. Holland, N. D., Holland, L. Z. & Holland, P. W. Scenarios for the making of vertebrates. *Nature* **520,** 450–455 (2015).
17. Hyman, L. H. *The invertebrates: smaller coelomate groups chaetognatha, hemichordata, pogonophora, phoronida, ectoprocta, brachipoda, sipunculida, the coelomate bilateria* Vol. 5. (McGraw-Hill, 1959).
18. Bourlat, S. J. *et al.* Deuterostome phylogeny reveals monophyletic chordates and the new phylum Xenoturbellida. *Nature* **444,** 85–88 (2006).
19. Philippe, H. *et al.* Acoelomorph flatworms are deuterostomes related to Xenoturbella. *Nature* **470,** 255–258 (2011).
20. Ruiz-Trillo, I., Riutort, M., Fourcade, H. M., Baguna, J. & Boore, J. L. Mitochondrial genome data support the basal position of Acoelomorpha and the polyphyly of the Platyhelminthes. *Mol. Phylogenet. Evol.* **33,** 321–332 (2004).
21. Hejnol, A. *et al.* Assessing the root of bilaterian animals with scalable phylogenomic methods. *Proc. R. Soc. Lond. B* **276,** 4261–4270 (2009).
22. Edgecombe, G. D. *et al.* Higher-level metazoan relationships: recent progress and remaining questions. *Org. Divers. Evol.* **11,** 151–172 (2011).
23. Srivastava, M., Mazza-Curll, K. L., van Wolfswinkel, J. C. & Reddien, P. W. Whole-body acoel regeneration is controlled by Wnt and Bmp-Admp signaling. *Curr. Biol.* **24,** 1107–1113 (2014).
24. Lartillot, N., Lepage, T. & Blanquart, S. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* **25,** 2286–2288 (2009).
25. Ulitsky, I., Shkumatava, A., Jan, C. H., Sive, H. & Bartel, D. P. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* **147,** 1537–1550 (2011).
26. Royo, J. L. *et al.* Transphyletic conservation of developmental regulatory state in animal evolution. *Proc. Natl Acad. Sci. USA* **108,** 14186–14191 (2011).
27. Putnam, N. H. *et al.* The amphioxus genome and the evolution of the chordate karyotype. *Nature* **453,** 1064–1071 (2008).
28. Irimia, M. *et al.* Extensive conservation of ancient microsynteny across metazoans due to cis-regulatory constraints. *Genome Res.* **22,** 2356–2367 (2012).
29. Sodergren, E. *et al.* The genome of the sea urchin *Strongylocentrotus purpuratus*. *Science* **314,** 941–952 (2006).
30. Freeman, R. *et al.* Identical genomic organization of two hemichordate hox clusters. *Curr. Biol.* **22,** 2053–2058 (2012).

31. Ikuta, T. *et al.* Identification of an intact ParaHox cluster with temporal colinearity but altered spatial colinearity in the hemichordate *Ptychodera flava*. *BMC Evol. Biol.* **13,** 129 (2013).
32. Cameron, R. A. *et al.* Unusual gene order and organization of the sea urchin hox cluster. *J. Exp. Zoolog. B Mol. Dev. Evol.* **306,** 45–58 (2006).
33. Baughman, K. W. *et al.* Genomic organization of Hox and ParaHox clusters in the echinoderm, *Acanthaster planci. Genesis* **52,** 952–958 (2014).
34. Santagati, F. *et al.* Identification of *cis*-regulatory elements in the mouse *Pax9/Nkx2–9* genomic region: implication for evolutionary conserved synteny. *Genetics* **165,** 235–242 (2003).
35. Lowe, C. J. *et al.* Dorsoventral patterning in hemichordates: insights into early chordate evolution. *PLoS Biol.* **4,** e291 (2006).
36. Wang, W., Zhong, J., Su, B., Zhou, Y. & Wang, Y. Q. Comparison of *Pax1/9* locus reveals 500-Myr-old syntenic block and evolutionary conserved noncoding regions. *Mol. Biol. Evol.* **24,** 784–791 (2007).
37. Santagati, F. *et al.* Comparative analysis of the genomic organization of *Pax9* and its conserved physical association with *Nkx2–9* in the human, mouse, and pufferfish genomes. *Mamm. Genome* **12,** 232–237 (2001).
38. Wang, S., Zhang, S., Zhao, B. & Lun, L. Up-regulation of C/EBP by thyroid hormones: a case demonstrating the vertebrate-like thyroid hormone signaling pathway in amphioxus. *Mol. Cell. Endocrinol.* **313,** 57–63 (2009).
39. Lowe, C. J. *et al.* Anteroposterior patterning in hemichordates and the origins of the chordate nervous system. *Cell* **113,** 853–865 (2003).
40. Kokubu, C. *et al.* A transposon-based chromosomal engineering method to survey a large cis-regulatory landscape in mice. *Nature Genet.* **41,** 946–952 (2009).
41. Giacopuzzi, E., Bresciani, R., Schauer, R., Monti, E. & Borsani, G. New insights on the sialidase protein family revealed by a phylogenetic analysis in metazoa. *PLoS ONE* **7,** e44193 (2012).
42. Harduin-Lepers, A., Mollicone, R., Delannoy, P. & Oriol, R. The animal sialyltransferases and sialyltransferase-related genes: a phylogenetic approach. *Glycobiology* **15,** 805–817 (2005).
43. Harduin-Lepers, A. *et al.* Evolutionary history of the alpha2,8-sialyltransferase (ST8Sia) gene family: tandem duplications in early deuterostomes explain most of the diversity found in the vertebrate ST8Sia genes. *BMC Evol. Biol.* **8,** 258 (2008).
44. Pardos, F. Fine structure and function of pharynx cilia in *Glossobalanus minutus* Kowalewsky (Enteropneusta). *Acta Zoologica* **69,** 1–12 (1988).
45. Kaul-Strehlow, S. & Stach, T. A detailed description of the development of the hemichordate *Saccoglossus kowalevskii* using SEM, TEM, Histology and 3D-reconstructions. *Front. Zool.* **10,** 53 (2013).
46. Ruppert, E. E., Cameron, C. B. & Frick, J. E. Endostyle-like features of the dorsal epibranchial ridge of an enteropneust and the hypothesis of dorsal-ventral axis inversion in chordates. *Invertebr. Biol.* **118,** 202–212 (1999).
47. Range, R. & Lepage, T. Maternal Oct1/2 is required for Nodal and Vg1/Univin expression during dorsal-ventral axis specification in the sea urchin embryo. *Dev. Biol.* **357,** 440–449 (2011).
48. Massagué, J. TGFβ signalling in context. *Nature Rev. Mol. Cell Biol.* **13,** 616–630 (2012).
49. Range, R. *et al. Cis*-regulatory analysis of nodal and maternal control of dorsal-ventral axis formation by Univin, a TGF-β related to Vg1. *Development* **134,** 3649–3664 (2007).

Supplementary Information is available in the online version of the paper.

Acknowledgements The *Ptychodera flava* genome project was supported by MEXT and OIST, Japan. This research was supported by USPHS grant HD42724 and NASA grant FDNAG2–1605 to J.G.; USPHS grant HD37277 to M.W.K.; NASA - NNX13AI68G to C.L. F.M. was funded by FP7/ERC grant [268513]. O.S. and D.S.R, and T.K. and N.S. were supported by the Molecular Genetics Unit and Marine Genomics Unit of the Okinawa Institute of Science and Technology Graduate University, respectively. Y.-H.S. and J.-K.Y. are supported by Academia Sinica and Ministry of Science and Technology, Taiwan. L.P. was supported by NIH grant R01HD073104. The *Saccoglossus kowalevskii* genome project was supported by a grant from the National Human Genome Research Institute, National Institutes of Health (U54 HG003273) to R.A.G.

Author Contributions J.Q., K.C.W. assembled the initial *S. kowalevskii* genomic assembly and performed quality assessments of the genome assemblies. *Ptychodera* collection, genome sequencing, and assembly: T.K., K.T., A.S., R.K., H.G., M.F., M.I.K., N.A., S.Y., A.F., T.H. *Rhabdopleura* collection: A.S. *Saccoglossus* RNA sequencing and analysis: R.M.F., M.W.K., R.C., C.L.K., S.L.L., M.H., S.R., D.M.M., K.C.W. Genome sequence production: A.C., Y.D., H.H.D., S.D., M.H., S.N.J., C.L.K., S.L.L., L.R.L., D.M., L.V.N., G.O., J.Sa., S.R., K.C.W., D.M.M., L.P., B.F., M.W.K. *Saccoglossus* sequence finishing: S.D., Y.D., D.M.M. Final *Saccoglossus* assembly: J.J., J.Sc. *Saccoglossus* gene modelling and validation: T.M., J.B., J.H.F., A.M.P., M.W. *Ptychodera* gene modelling and analyses: T.K., R.K., K.H., E.S., F.G., K.W.B., K.T., O.S., J.G., N.S. Gene family analyses: O.S., T.K., F.M., L.P., R.M.F., C.L., J.G. Synteny: N.H.P., O.S., J.-X.Y. Repeats: O.S. *Saccoglossus* sequencing and assembly project management: S.R., D.M.M., K.C.W., R.A.G. *Ptychodera* expression analysis: Y.-C.C., Y.-H.S., J.-K.Y. Phylogenetic analyses: F.M. Additional EST collections: T.H.-K., K.T., A.S., A.T.S., J.P., P.G., C.C., C.L. HGT and novelties: J.G., O.S. Pharyngeal cluster analysis and expression: J.G., O.S., N.S., K.B., A.G. Project coordination, manuscript writing: O.S., T.K., F.M., K.T., N.S., J.G., C.L., D.S.R.

Author Information Sequencing data have been deposited in NCBI BioProject under accession number PRJNA12887 (*Saccoglossus kowalevskii*) and DDBJ under accession number PRJDB3182 (*Ptychodera flava*). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to O.S. (oleg.simakov@oist.jp), J.G. (jgerhart@berkeley.edu), N.S. (norisky@oist.jp) and D.S.R. (dsrokhsar@gmail.com).

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Sequencing. Sperm DNA from adult males was extracted for sequencing as described in Supplementary Note 2. A single male was used for each species to minimize the impact of heterozygosity on assembly. For *Saccoglossus*, approximately eightfold redundant random shotgun coverage (totalling 8.1 Gb) was obtained with Sanger dideoxy sequencing at the Baylor College of Medicine Genome Center, including 34,279 BAC ends and 459,052 fosmid ends. For *Ptychodera*, 1.3 Gb in Sanger shotgun sequences, 15.3 Gb in Roche 454 pyrosequence reads, and 52-Gb paired-end sequences with Illumina MiSeq, along with mate-pairs, were generated at the Okinawa Institute of Science and Technology Graduate University. More sequencing details are available in Supplementary Note 2.

Genome assemblies. We assembled the *Saccoglossus* genome with Arachne[50], combined with BAC/fosmid pair information to produce the final assembly. This *Saccoglossus* assembly includes 7,282 total scaffold sequences spanning a total length of 758 Mb. The relatively modest nucleotide heterozygosity (0.5%) of *S. kowalevskii*, coupled with longer read lengths, enabled assembly of a single composite reference sequence. Half of the assembly is in scaffolds longer than 552 kb (the N50 scaffold length), and 82% of the assembled sequence is found in 1,602 scaffolds longer than 100 kb. For *Ptychodera* we used the Platanus[51] assembler. The resulting total scaffold length was 1,229 Mb, with half the assembly in scaffolds longer than 196 kb (N50 scaffold length). *P. flava* exhibited a notably higher heterozygosity (1.3% single nucleotide heterozygosity with frequent indels) than *S. kowalevskii*, presumably related to its pelagic dispersal and larger effective population size[52]. We therefore initially produced stringent separate assemblies of the two divergent haplotypes, and found that many scaffolds had a closely related second scaffold with ~94% BLASTN identity (over longer stretches, including indels). To avoid reporting both haplotypes at these loci, scaffolds with less than 6% divergence over at least 75% of their length were merged into a single haploid reference for comparative analysis. To further classify regions with 'double' depth and single haplotype regions we implemented a Hidden Markov Model classifier. We find that at least 63% of the initial Platanus assembly constitutes merged haplotypes. The inferred SNP rate for those regions is 1.3%, while for the remaining haplotype regions it is below 0.1%. Further details of assemblies are described in Supplementary Note 2.

Gene predictions. Transcriptome data for both species were used, along with homology-guided and *ab initio* methods, to predict protein-coding genes (Supplementary Note 3). For *Saccoglossus*, 8.6 million RNAseq reads were generated from 7 adult tissues and 15 developmental stages using Roche 454 sequencing, along with previously deposited ESTs in GenBank. For *Ptychodera*, extensive EST data from egg, blastulae, gastrulae, larvae, juveniles, adult proboscis, stomochord, and gills defining 34,159 cDNA clones[53], and 879,000 Roche/454 RNAseq reads from a mixed library of developmental stages[54] were used. The *Saccoglossus* genome was annotated using JGI gene prediction pipeline[55], while Augustus[56] was used to produce gene models for *Ptychodera*. We find a total of 34,239 gene predictions for *Saccoglossus* (68% with transcript evidence) and 34,687 for *Ptychodera* (43% with transcript evidence), although these are overestimates of the true gene number due to fragmented gene predictions, mis-annotated repetitive sequences, and spurious predictions. As described in the main text, 18–19,000 gene models in each species have known annotations and/or orthologues in other species.

Gene family analysis. Gene family clustering was done using a progressive (leaf to root) BLASTP-based clustering algorithm, where at a given phylogenetic node the gene families are constructed taking into account protein similarities among ingroups and outgroups[57]. For the inference of deuterostome gene families we use the bilaterian node of the clustering. To call gene families present in the deuterostome ancestor, we required (1) at least two ambulacrarian orthologues out of the three available ambulacrarian genomes and at least two chordate orthologues, or (2) at least two deuterostomes (chordates and/or ambulacrarians) and two outgroups in the bilaterian level clusters.

Transposable elements. Repetitive sequences were identified using RepeatScout[58], followed by manual curation and annotation using both a Repbase release (version 20140131)[59] and BLASTX-based search against a custom collection of transposons, using a previously described repeat identification and annotation pipeline[57] (Supplementary Note 5). The assemblies were then masked with RepeatMasker version open-4.0.5[60]. The repetitive complements of the two hemichordate genomes are summarized in Supplementary Table 5.1.

Phylogenetic analysis. Phylogenetic analyses were done using metazoan-level gene family clusters based on whole-genome sequences (Supplementary Note 4), selecting a single orthologue per genome with the best cumulative BLASTP to other species, and best reciprocal BLASTP hits to species with transcriptome-only information (Supplementary Note 6). Single gene alignments were built using Muscle[61] and filtered using Trimal[62] for each orthologue, and were concatenated, yielding a supermatrix of 506,428 positions with 34.9% missing data. This supermatrix was analysed with ExaML assuming a site-homogenous LG+Γ_4 model partitioned for each gene[63]. A slow-fast analysis was conducted to stratify marker genes based on the length of the branch leading to acoels in individual trees. A subset of the slowest 10% of genes was analysed with the site-heterogenous CAT+GTR+Γ_4 model using Phylobayes[24]. Molecular dating was carried out using Phylobayes[24] using the log-normal relaxed clock model and the calibrations described in Supplementary Table 6.2.

Synteny analysis. Macro- and micro-syntenic linkages were calculated as described in Supplementary Note 7. For Fig. 3a, we merged the amphioxus scaffolds into 17 pre-defined scaffold groups as suggested in ref. 27. These 17 merged scaffold groups represent the 17 ancestral linkage groups (ALGs) shared in chordates. Then we calculated the orthologous gene groups shared by each amphioxus ALG–*Saccoglossus* scaffold pair and generated the dot plot as described in Supplementary Note 7. For micro-synteny we required at least three genes (separated by a maximum of ten genes) to be present in pairwise comparisons. Under random reshuffling of the genome, this yields 10% false positives in pairwise genome comparisons, that is, we observe approximately one-tenth as many micro-syntenic blocks between the two genomes when gene orders are shuffled. This false-positive rate, however, falls to 1% when considering more than two species. For our inference of deuterostome ancestral and novel synteny we therefore focus on blocks present in at least three species (and both ingroup representatives, that is, ambulacrarians and chordates). This yields 698 blocks that can be traced back to the deuterostome ancestor, including 71 blocks found exclusively in deuterostome species (shared among ambulacrarians and chordates), including the pharyngeal cluster discussed in Fig. 4.

Whole-genome alignment. Whole-genome alignments were conducted with MEGABLAST[64] using parameters previously reported[65]. We assessed the distribution of the resulting 12,722 aligned loci across known gene annotations in ENSEMBL[66], previously identified conserved pan-vertebrate elements[65], as well as known enhancers in human according to LBL database[67].

Gene novelties. Deuterostome gene novelties were assessed initially through bilaterian gene clusters (Supplementary Note 10) by requiring at least two species on both ambulacrarian and chordate side to be present. The novelties were further automatically subdivided into four categories: G1 (gain type I), with no BLASTP hit outside of deuterostomes; G2 (gain type II), with a novel PFAM domain present only in deuterostomes; G3 (gain type III) having a novel PFAM combination unique to deuterostomes; and G4 (gain type IV), those that do not fall under any of the G1–3 categories and define novelties due to acceleration in the substitution rate on the deuterostome stem. To confirm the novel nature, especially for G4 novelties, we have constructed phylogenies for the members and non-deuterostome BLASTP hits (up to an *e*-value of 1×10^{-20}) using MAFFT-alignment-based FastTree calculations. The trees were assessed for the accelerated rate of evolution at the deuterostome stem (Supplementary Fig. 9.1.1). The final result is provided in the Supplementary Information.

Curation of candidates for horizontal gene transfer on the deuterostome stem. We examined in detail gene families found broadly in deuterostomes whose encoded peptides were readily alignable to microbial sequences but had no detectable similarity in non-deuterostome animals. Criteria for evaluation included: (1) the hemichordate gene matches microbial genes at least ten orders of magnitude in the *e*-value better than it matches sequences of non-deuterostome metazoans (most of the putative HGTs we describe have no non-deuterostome metazoan hit at all); (2) it has a defined genomic locus among bona fide metazoan genes; (3) it shares an exon–intron structure with genes of chordates and other ambulacraria; and (4) when a low bitscore match is found to a non-deuterostome metazoan sequence, that sequence is identified as containing different domains (domain structure according to CDD[68]) and/or different exon–intron structure, implying dubious relatedness. When phylogenetic trees are constructed for these HGT-candidate proteins, the trees contain numerous branches for microbial sequences and none for non-deuterostome metazoan sequences, or only very long branches for dubiously relatives, and hence the trees differ greatly from the metazoan species tree, except within the deuterostome clade.

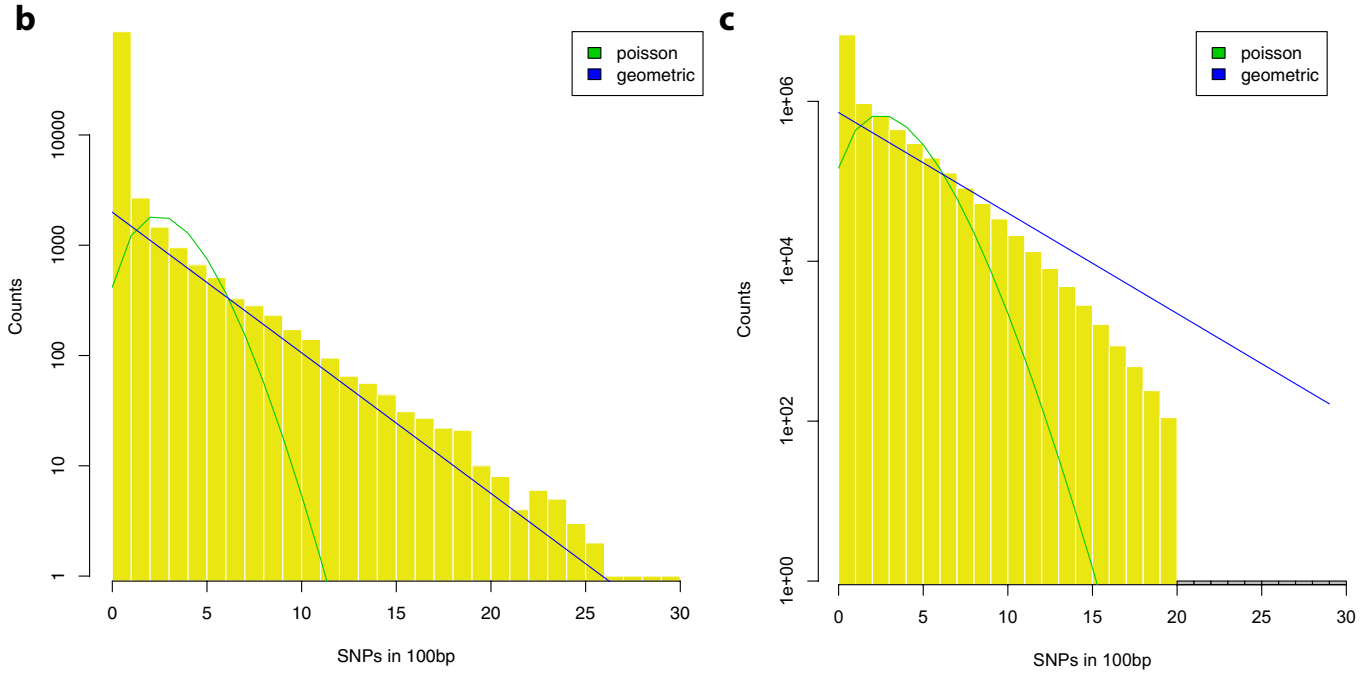
Code availability. Original data and code can be accessed at https://groups.oist.jp/molgenu.

50. Jaffe, D. B. *et al.* Whole-genome sequence assembly for mammalian genomes: Arachne 2. *Genome Res.* **13,** 91–96 (2003).
51. Kajitani, R. *et al.* Efficient *de novo* assembly of highly heterozygous genomes from whole-genome shotgun short reads. *Genome Res.* **24,** 1384–1395 (2014).
52. Romiguier, J. *et al.* Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature* **515,** 261–263 (2014).

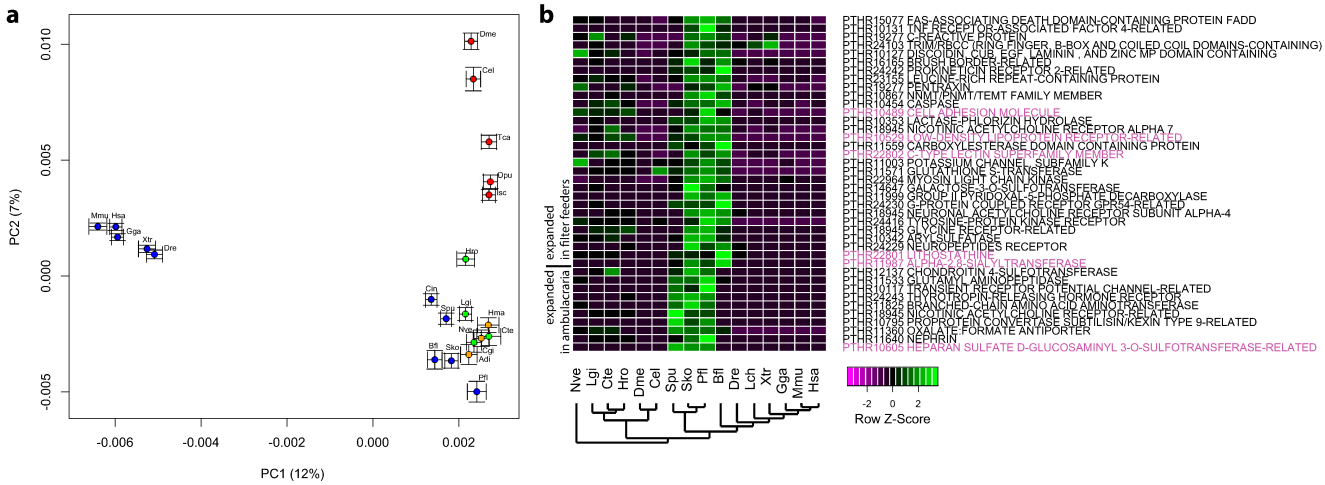
53. Tagawa, K. *et al.* A cDNA resource for gene expression studies of a hemichordate, *Ptychodera flava*. *Zoolog. Sci.* **31,** 414–420 (2014).
54. Chen, S. H. *et al.* Sequencing and analysis of the transcriptome of the acorn worm *Ptychodera flava*, an indirect developing hemichordate. *Mar. Genomics* **15,** 35–43 (2014).
55. Salamov, A. A. & Solovyev, V. V. *Ab initio* gene finding in *Drosophila* genomic DNA. *Genome Res.* **10,** 516–522 (2000).
56. Stanke, M. & Waack, S. Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics* **19** (Suppl. 2), ii215–ii225 (2003).
57. Simakov, O. *et al.* Insights into bilaterian evolution from three spiralian genomes. *Nature* (2013).
58. Price, A. L., Jones, N. C. & Pevzner, P. A. *De novo* identification of repeat families in large genomes. *Bioinformatics* **21** (Suppl. 1), i351–i358 (2005).
59. Jurka, J. *et al.* Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet. Genome Res.* **110,** 462–467 (2005).
60. Smit, A., Hubley, R. & Green, P. RepeatMasker http://www.repeatmasker.org. (2007).
61. Edgar, R. C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5,** 113 (2004).
62. Capella-Gutiérrez, S., Silla-Martinez, J. M. & Gabaldon, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25,** 1972–1973 (2009).
63. Aberer, A. & Stamatakis, A. ExaML: Exascale maximum likelihood: program and documentation. See http://sco.h-its.org/exelixis/web/software/examl/index.html (2013).
64. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25,** 3389–3402 (1997).
65. Lee, A. P., Kerk, S. Y., Tan, Y. Y., Brenner, S. & Venkatesh, B. Ancient vertebrate conserved noncoding elements have been evolving rapidly in teleost fishes. *Mol. Biol. Evol.* **28,** 1205–1215 (2011).
66. Cunningham, F. *et al.* Ensembl 2015. *Nucleic Acids Res.* **43,** D662–D669 (2015).
67. Visel, A., Minovitsky, S., Dubchak, I. & Pennacchio, L. A. VISTA Enhancer Browser–a database of tissue-specific human enhancers. *Nucleic Acids Res.* **35,** D88–D92 (2007).
68. Marchler-Bauer, A. *et al.* CDD: NCBI's conserved domain database. *Nucleic Acids Res.* **43,** D222–D226 (2015).
69. Marinić, M., Aktas, T., Ruf, S. & Spitz, F. An integrated holo-enhancer unit defines tissue and gene specificity of the Fgf8 regulatory landscape. *Dev. Cell* **24,** 530–542 (2013).

a

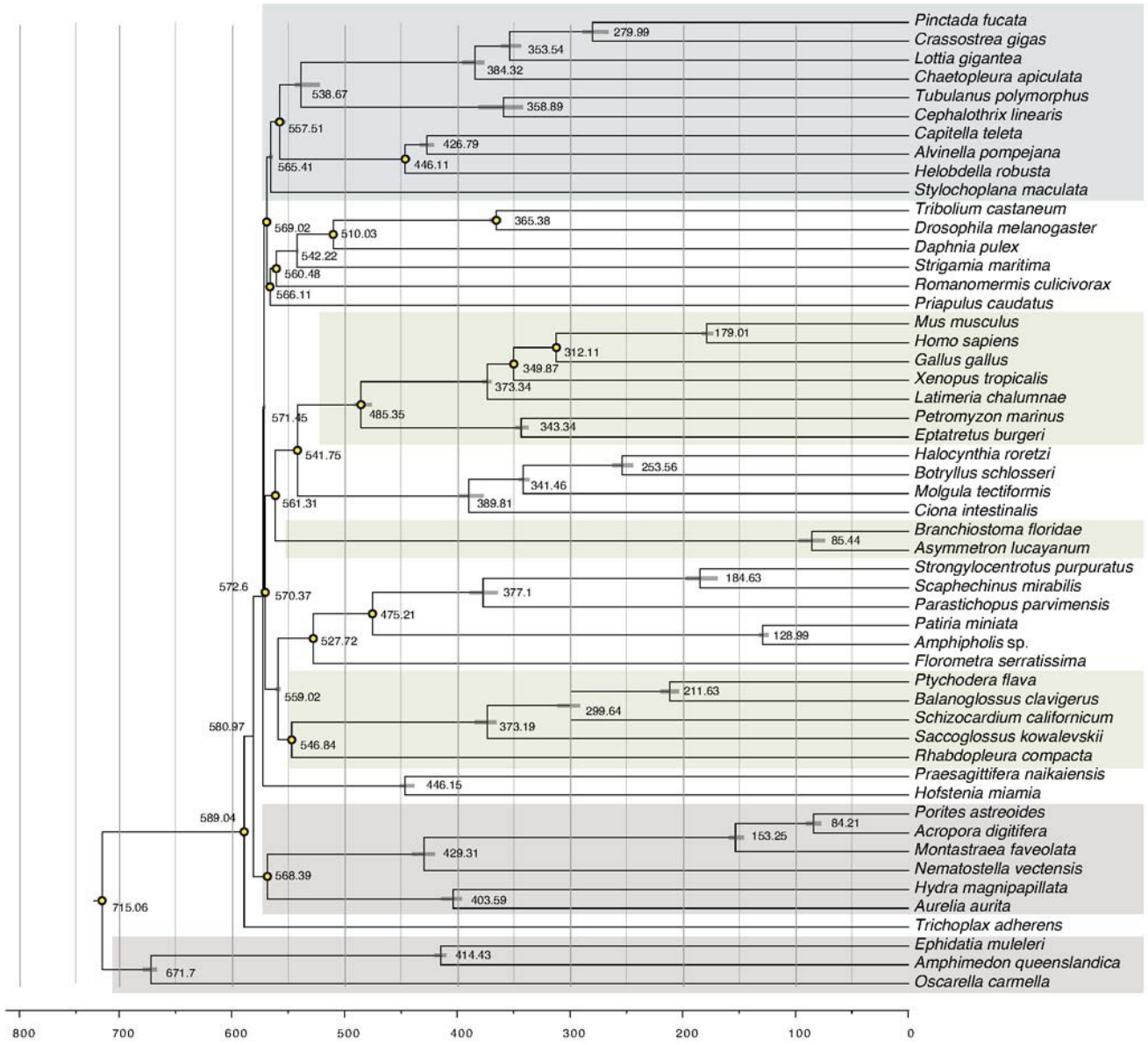
	Saccoglossus	*Ptychodera*
Scaffold total	7,282	218,255
Contig total	20,913	322,077
Scaffold sequence total, Mb	758	1,229
Scaffold N50, kb	552	196
Gene models	34,239	34,687
SNP rate, %	0.5	1.3 (2x), 0.06 (1x)



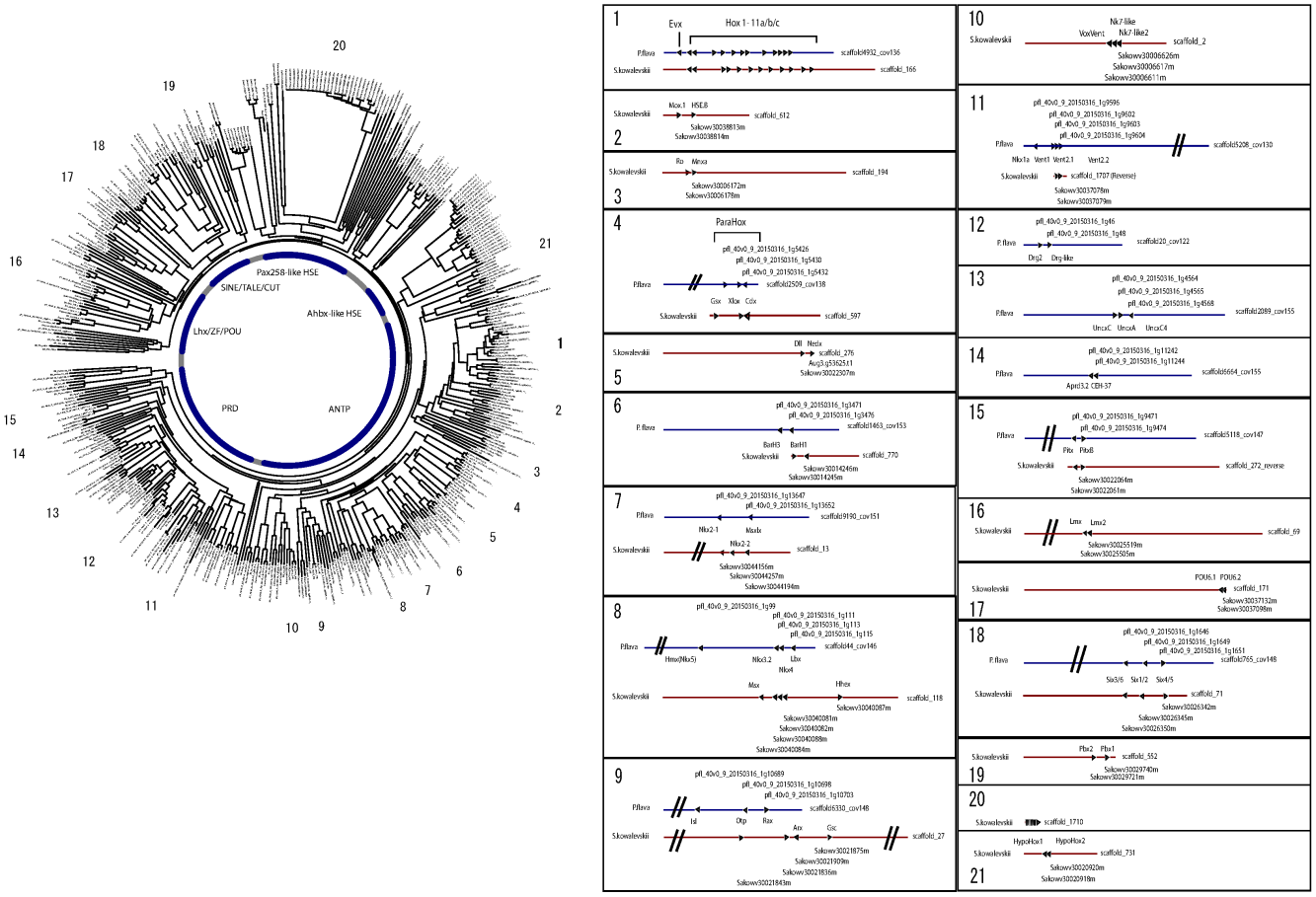
Extended Data Figure 1 | Summary of genome assemblies and heterozygosity distributions for *Saccoglossus* and *Ptychodera*. **a**, Genome statistics summary. **b**, **c**, The single nucleotide polymorphism distribution across 100-bp windows for *Saccoglossus* (**b**) and the corresponding distribution for *Ptychodera* (**c**). The distributions in **b** and **c** are fitted with a geometric (expected when high recombination rate is present) and a Poisson distribution (expected with low recombination rate). The distribution for *Saccoglossus* is fitted to windows with one or more SNPs only, as there is an excess of zero SNP windows (approximately 84% of total 94,324 selected windows). For methods refer to Supplementary Note 2.



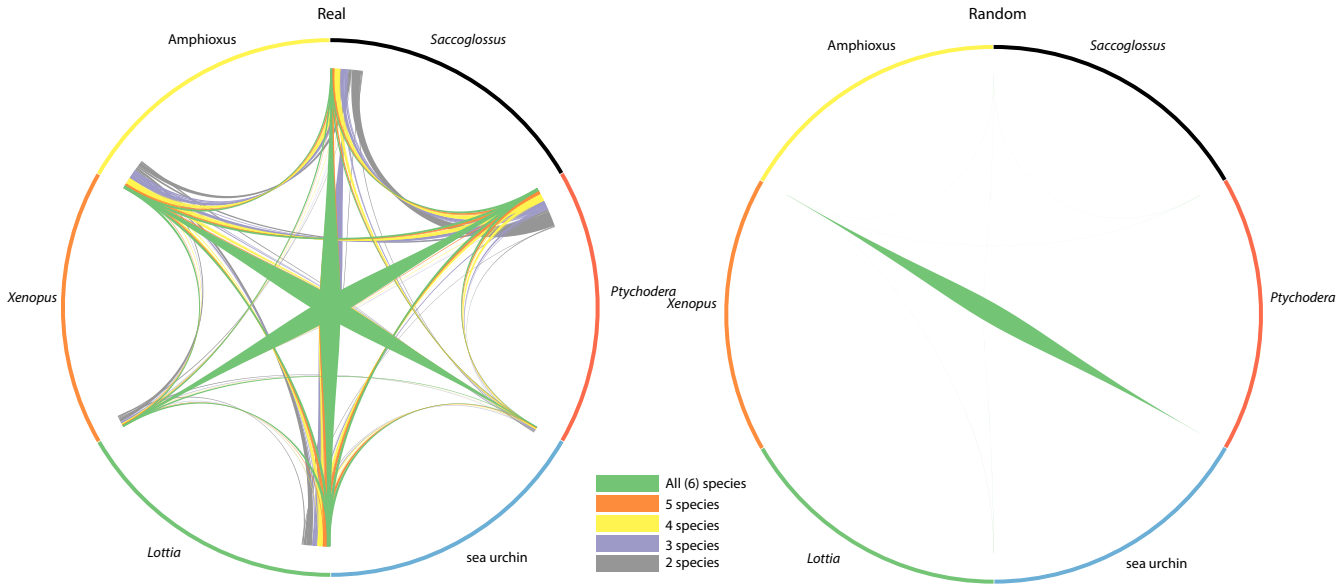
Extended Data Figure 2 | Ambulacrarians approximate the ancestral metazoan gene repertoire. a, Principal component analysis of Panther gene family sizes. Variances of the first two components are plotted in parentheses. Blue indicates deuterostomes; green indicates lophotrochozoans; red, ecdysozoans; yellow/orange, non-bilaterian metazoans. Note the clustering of the ambulacrarians Sko, Pfl and Spu with the non-vertebrate deuterostomes Bfl and Cin in the lower right corner, also with the lophotrochozoans Cgi, Lgi, Hro, Cte and the non-bilaterians Hma, Nve and Adi. **b**, Heat map of gene family counts showing significant (Fisher's exact test P value <0.01 after Bonferroni multiple testing correction) expansion in ambulacrarians as well as in *Saccoglossus*/*Ptychodera*/amphioxus. The cases discussed in the main text are highlighted in red. See Supplementary Note 4 for details. Species abbreviations are defined in Supplementary Note 4.1.



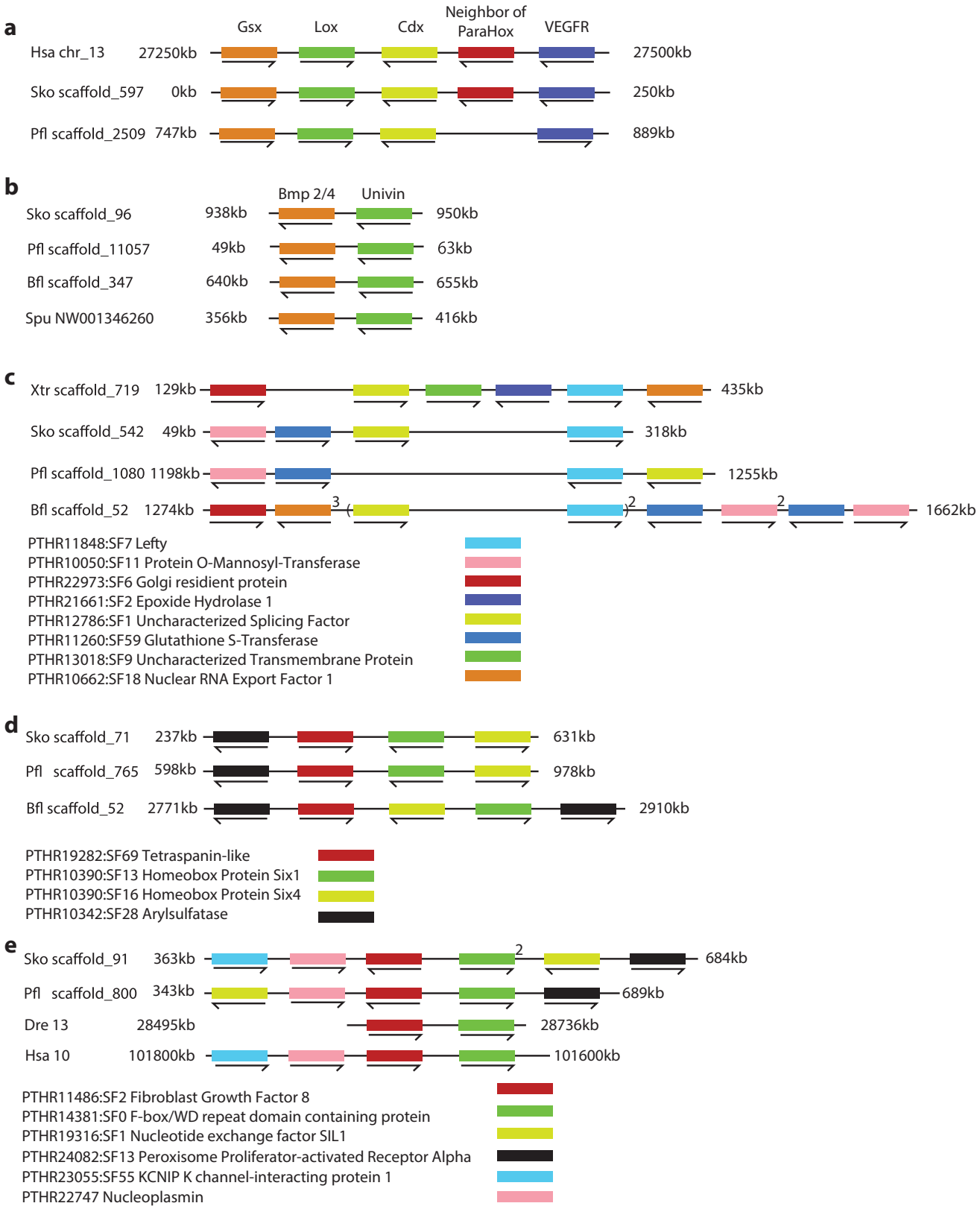
Extended Data Figure 3 | Molecular dating of deuterostome and metazoan radiations using PhyloBayes assuming a log-normal relaxed clock model. Yellow circles on particular nodes indicate the calibration dates applied from the fossil record, as indicated in Supplementary Note 6.2. Bars are 95% credibility intervals derived from posterior distributions. Note the estimated times of divergence of chordates and ambulacraria (the deuterostome ancestor) at 570 million years ago (Ma; mid-Ediacaran), hemichordates and echinoderms at 559 Ma, enteropneusts and pterobranchs at 547 Ma, and Harrimaniid and Ptychoderid enteropneusts at 373 Ma.



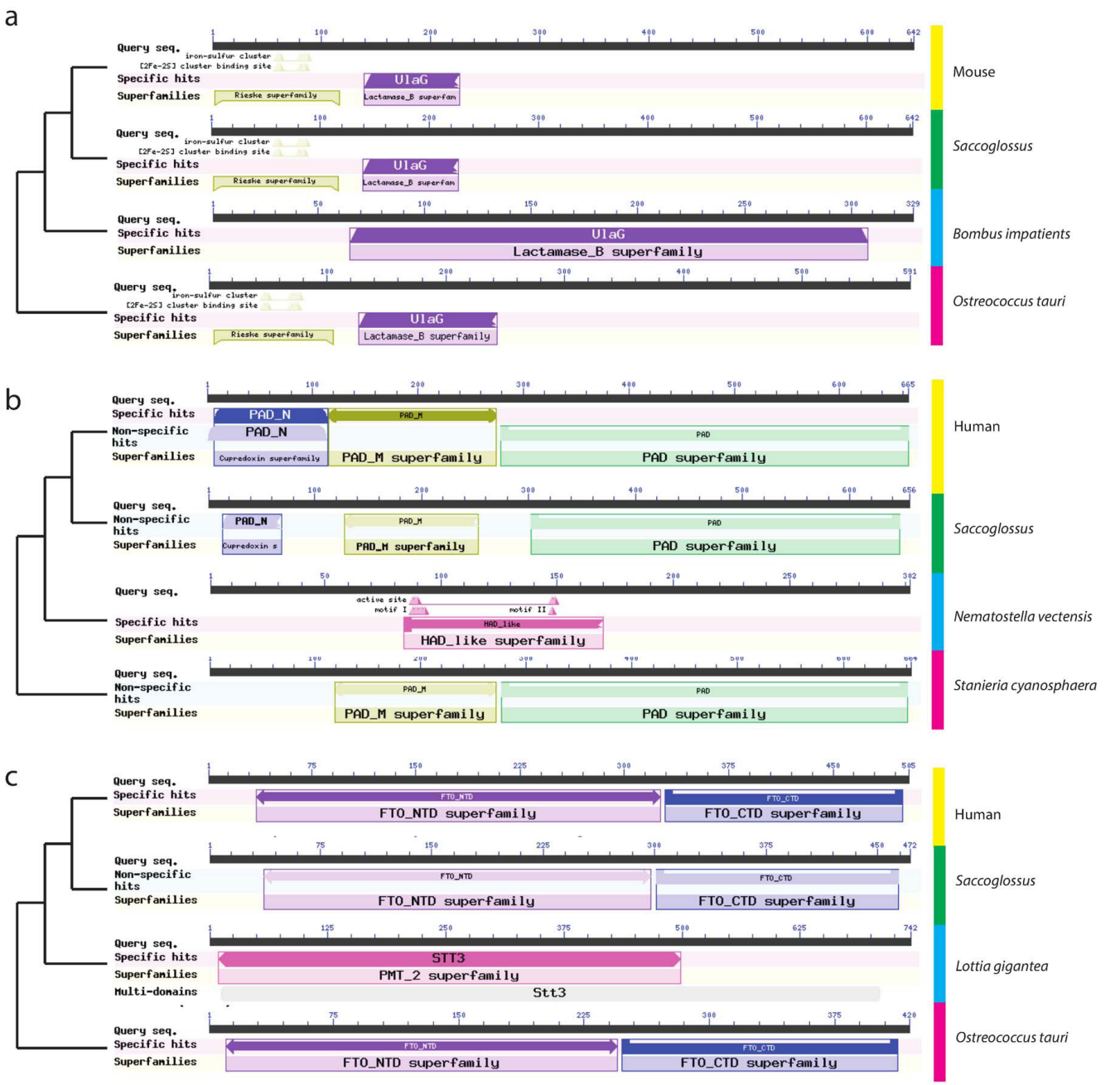
Extended Data Figure 4 | Homeobox gene complement of the two hemichordates in comparison to that of amphioxus. The numbers of homeobox-containing gene models are 170 in *Saccoglossus* and 139 in *Ptychodera*. These homeobox domains were aligned with 128 homeobox genes of *Branchiostoma floridae* using ClustalW2, then gaps and unaligned regions were manually removed. Since some genes have more than one homeobox domain, we kept all domains or chose the longest one according to the state of domain conservation. In total, 448 homeobox sequences were aligned. See Supplementary Information for details. The clusters of homeobox genes on scaffolds in *Saccoglossus* and *Ptychodera* were identified and drawn at positions around the tree. Conserved clusters between the two species were aligned. In addition to the well-known Hox and ParaHox cluster, 17 clusters were found in at least one of the hemichordates or some in both. Sixteen genes of the Nkx class are distributed over four clusters: (i) *nkx1a-vent1-vent2.1-vent2.2*; (ii) *nkx2.1-nkx2.2-msxlx*; (iii) *nkx5-msx-nkx3.2-nkx4-lbx-hex*; and (iv) *voxvent-nk7like-nk7like2*. The second cluster (ii) of these is part of the pharyngeal cluster (Fig. 4). Another five-gene cluster consists of one Lim class homeobox gene and four PRD class homeobox genes; *isl-otp-rax-arx-gsc*. A cluster of *six3/6-six1/2-six4/5* was found in both species, and a cluster of three *unx* genes was found only in *P. flava*. Ten more clusters were found containing two homeobox genes each. Notably, we found species-specific homeobox clusters in both species. Three remarkable clusters were found in *S. kowalevskii* in which 10, 12 and 5 homeobox-containing genes are tandem duplicated in scaffold_1710, _52 and _4796, respectively. We also found such clusters in *P.flava* in which 7, 4, 8 and 10 genes are aligned on scaffold 19451, scaffold 1398, scaffold 12422 and scaffold 154657, respectively. All homeobox genes identified in the genomes of the two hemichordates and amphioxus are listed in the Supplementary Table for Extended Data Fig. 4. This list includes some genes not containing a homeobox (for example, *pax1/9*) in cases where other family members do (for example, *pax2*).



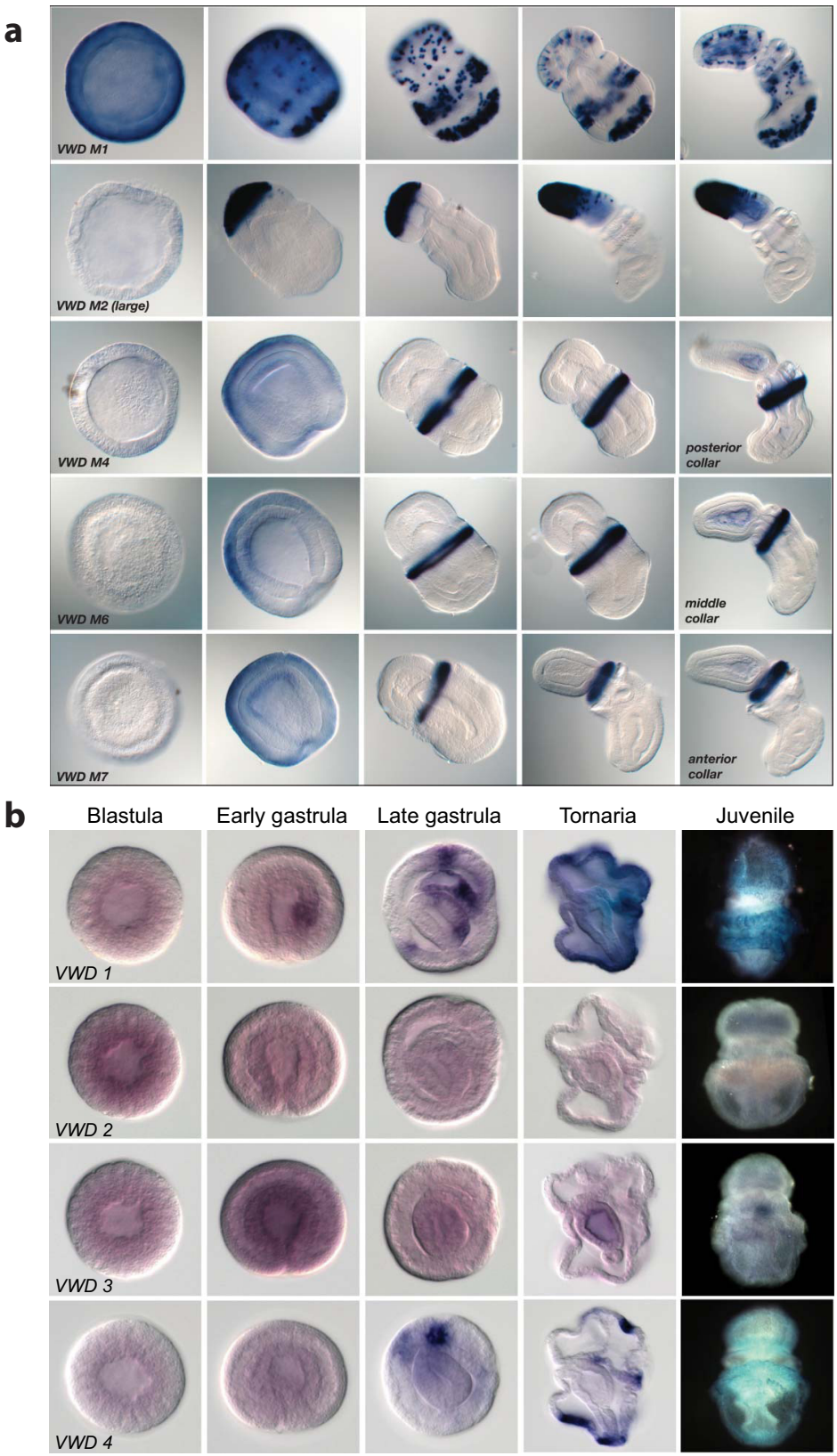
Extended Data Figure 5 | High retention rate of micro-synteny in *Saccoglossus*. Circos plot showing micro-syntenic conservation in blocks of genes ($n_{max} = 10$ and $n_{min} = 2$) for six metazoan species for observed (left) and simulated (right) linkages. The width of connecting segments is proportional to the number of genes participating in the syntenic linkages (normalized by the total gene count). In this representation scaffolds are placed end-to-end, and adjacent scaffolds need not be from the same chromosome. While simulated data yields some blocks shared between pairs of species, few or no synteny blocks can be recovered among three or more species (Methods). *Saccoglossus* shows one of the highest retentions among the selected species (and the highest among the sequenced ambulacrarians). *Xenopus* (and vertebrates in general) have lost some micro-synteny due to whole-genome duplications and differential loss of paralogues. The matching between the hemichordate *S. kowalevskii* and the chordate amphioxus is highest, consistent with the fact that neither genome has undergone extensive gene loss (as have tunicates) or pseudo-tetraploidization with extensive loss of paralogues (as have vertebrates).



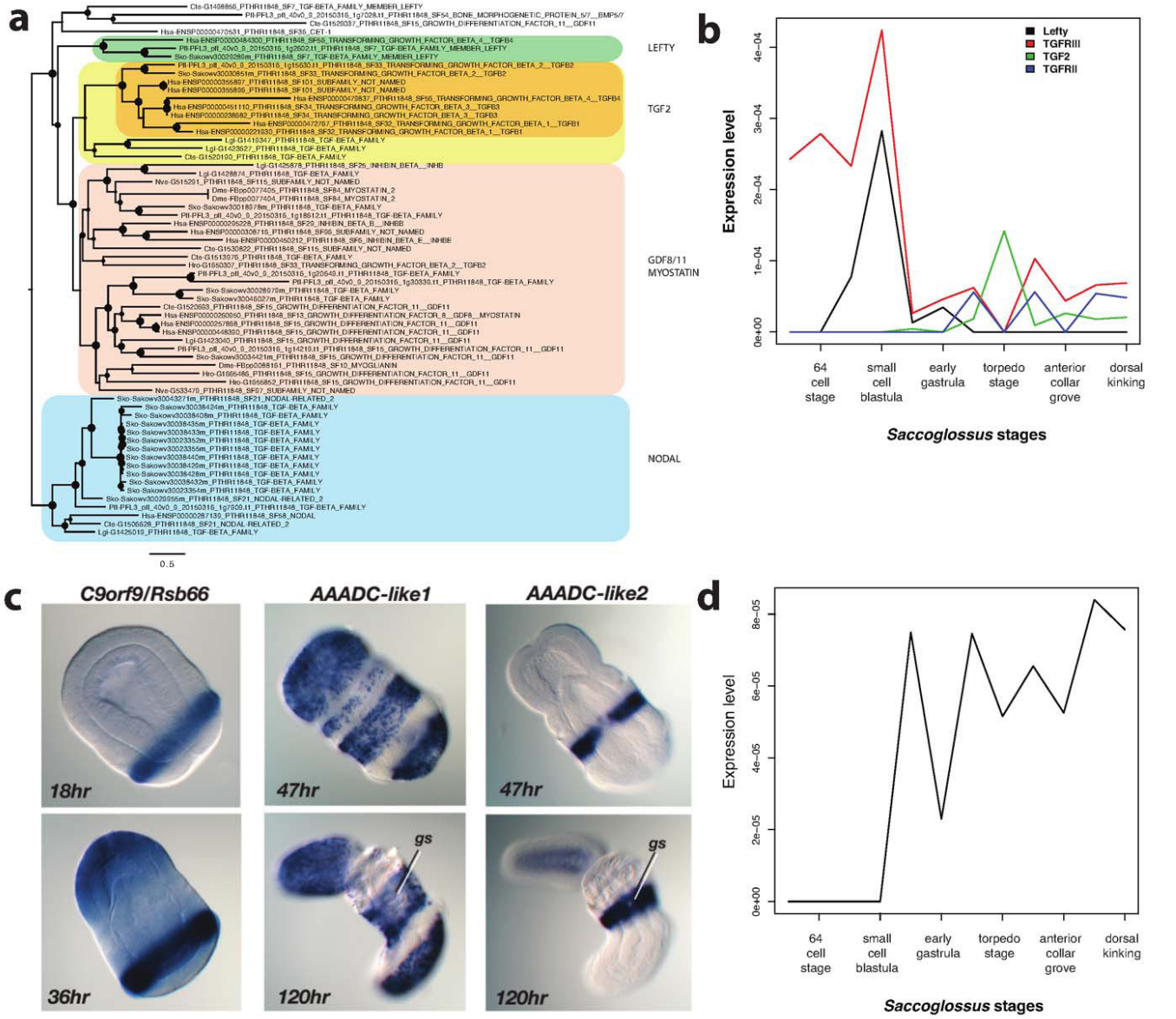
Extended Data Figure 6 | Deuterostome specific micro-syntenic linkages. a, **b**, Very tight linkages with no intervening genes. **a**, ParaHox cluster shown in *S. kowalevskii*, *P. flava*, and human. **b**, *bmp2/4* and *univin* cluster in the hemichordates *S. kowalevskii and P. flava*, the sea urchin *S. purpuratus*, and the cephalochordate *B. floridae*. **c**–**e**, Loose micro-syntenic linkages with a maximum of five intervening genes: *lefty* (**c**), *six1–six4* (**d**), and *fgf8–fbxw* (**e**)[69] clusters. For **c** to **e** all species with micro-synteny are shown. Numbers above the genes indicate the copy number in the locus.



Extended Data Figure 7 | Three examples showing the domain structures of some proteins encoded by genes found in deuterostomes and marine microbes but not non-deuterostome animals. Best BLASTP hits of the *Saccoglossus* sequence in human/mouse, as well as in non-deuterostome metazoans and in non-metazoans (such as the cyanobacterium *Staniera cyanosphaera*, or the eukaryotic micro-alga *Ostreococcus tauri*) are shown. **a**, Cytidine monophosphate-*N*-acetylneuraminic acid hydroxylase (CMAH), an enzyme of sialic acid modification; **b**, peptidyl arginyl deiminase (PAD), an enzyme of post-translational modification of proteins; **c**, FATSO-like, also called α-ketoglutarate-dependent dioxygenase FTO, an enzyme that de-methylates N^6-methyladenosine in nuclear RNA. Other analyses of these and other genes with the unusual phylogenetic distributions can be found in Supplementary Note 10.



Extended Data Figure 8 | *In situ* hybridization demonstration of the expression of von Willebrand type D (vWD) domain-encoding genes (putative glycoproteins/mucins) in *Saccoglossus* and *Ptychodera*. **a**, In *Saccoglossus* the genes are specifically expressed in different subregions of the ectoderm of the proboscis or collar at these pre-feeding stages. **b**, In *Ptychodera*, several of the genes are expressed in endoderm as well as ectoderm of the developing tornaria larva. The sequence IDs for the genes are provided in Supplementary Note S10.4.



Extended Data Figure 9 | Gene innovation in deuterostomes. a, FastTree phylogenetic tree of the TGFβ family members Lefty, TGFβ2, GDF8/11 and Nodal ligands (using GTR model). Bootstrap support is plotted as filled circles (size proportional to the support value) on each node. While Lefty shows deuterostome unique sequence composition, TGFβ2 has an acceleration of sequence change at the deuterostome stem branch, compared to the GDF8/11 or Nodal groups. **b**, Temporal co-expression of Lefty and TGFβ receptor type III in *Saccoglossus* at pre-gastrulation developmental stages and of TGFβ2 and TGFβ receptor type II at post-gastrulation stages. **c**, *In situ* hybridization demonstration of the expression in *S. kowalevskii* of one of the putative type I novelty genes (*c9orf9*, also known as *rsb66*) and of two of AAADC genes (aromatic amino acid decarboxylases of the microbial type) of *S. kowalevskii* (also in *P. flava* and *B. floridae*), which closely resemble sequences from bacteria rather than from non-deuterostome metazoans. gs, gill slits. **d**, The temporal expression profile for *c9orf9* during *S. kowalevskii* development, taken from transcriptome data.

Extended Data Table 1 | Examples of deuterostome gene novelties and their genomic features

ID	gene family name	BLAST e-value in non-deuterostome metazoa? # or "none"	BLAST e-value in non-metazoan clades with these domains? E.g., *Ostreococcus, Micromonas*, etc.	Phylogeny support for novelty/HGT	Origin of novelty?	Putative function in deuterostomes?
1	Lefty	*e-12, partial prodomain but no ligand domain in Capitella*	-	EDF9	type IV	antagonist of Nodal signaling. Developmental patterning?
2	Univin, Vg1, DVR, GDF1	*weak matches to bmp2/4 sequences*	-	Figure S10.3.1	type IV	agonist of Nodal signaling. Developmental patterning?
3	TGFb2	*e-35 for cnidarian Aiptasia and sponge Sycon, no protosotmes*	-	EDF9	type IV	Signaling via Smad2/3 activation. Regulation of cell activities
4	Thrombospondin1/2	*none, except the sponge Amphimedon has a partial combination of TSP1 and TSP3 domains*	-	unique domain combination due to introduction of 3 TSP1 domains	type II	Activation of TGFb2 signals
5	TGFbR2	no matches to deuterostome ectodomain; many matches to deuterostome protein kinase domain	-	Figure S10.3.3	type IV	Receptor specific for TGFb2
6	UDP-N-acetylglucosamine 2-epimerase/N-acetylmannosamine kinase	*Symsagittifera (acoel) e-154 matching both domains, but not in Hofstenia (acoel); no protostome match*	*e-94 to e-85 for the epimerase domain in Parcubacterium (bacteria), Prochlorococcus (cyanobacterium); Micromonas (green microalga). e-44 to e-30 for the kinase domain in Mesorhizobium and other bacteria*	No non-deuterostome metazoan sequenes in gene phylogeny, except acoels of uncertain relatedness to deuterostomes	HGT/convergence	first step of deuterostome sialic acid synthesis; different from protostome first step
7	CMP-N-Acetylneuraminic acid hydroxylase (modifies sialic acid)	*e-2 for unrelated match in Caenorhabditis brenneri, neither domain*	*e-127 to e-86, both domains, Ostreococcus and Micromonas (green microalga)*	No non-deuterostome metazoan sequences in gene phylogeny (S10)	HGT/convergence	produces a glycolyl sialic acid that is also added to glycoconjugates in deuterostomes
8	alpha-2,6-GalNac-sialyltransferase (expanded family in Amphioxus and ambulacraria)	*e-14 match to ST6Gal sialyltransferases of Tribolium, Drosophila, and other insects, with different specificity for oligosaccharide terminus.*	*e-10 to e-8 for Bathycossus (green microalga) and Emeliania (coccolithophore)*	unclear; deuterostome ST6GalNAc connected weakly to ST6Gal (bilaterian ancestral sequence)	type IV	New specificity of sialic linkage compared to protostomes
9	alpha-2,3-sialyltransferases (expanded family in Amphioxus and ambulacraria)	*e-61 match to Oscarella (sponge) ST3 sequence., and then e-7 match to ST6Gal of insects, with different specificity for oligosacharide terminus*	*e-17 to e-10 matches to Bathycossus ST3s (green microalga) and Emeliania (coccolithophore)*	unclear; deuterostome ST3 connected to sponge ST3 and weakly to ST6Gal (bilaterian ancestral sequence)	type IV	New specificity of sialic linkage compared to protostomes
10	alpha-2,8-sialyltransferases (greatly expanded family in Amphioxus and ambulacraria)	*e-7 to ST6Gal of insects, with different specificity for oligosacharide terminus*	*e-12 to e-8 matches to Bathycossus ST8s (green microalga) and Emeliania (coccolithophore)*	unclear; deuterostome ST8 connected weakly to sponge ST3 and ST6Gal (bilaterian ancestral sequence)	type IV	New specificity of sialic linkage compared to protostomes
11	B4GALNT, adds N-AcetylGal to sialic acid containing oligosaccharides (expanded family in Amphioxus and ambulacraria)	none	*Nitratifractor (epsilon- proteobacterium)*	No non-deuterostome metazoans in gene phylogeny	HGT/convergence	synthesis of gangliosides, not found in protostomes
12	sialidase 1,2,3,4 (expanded family in Amphioxus and ambulacraria)	*e-77 for NEU1 in Nematostella vectensis (shares an iontron boundary); e-40 for NEU1 of the sponge Oscarella. Weak match (>e-20) of NEU2,3,4 to Nematostella and sponge; no protostome matches*	*e-68 to e-60 to NEU1-like matches in Blastopirellula, Sphingobacterium and many other bacteria. e-50-e-40 to Neu5-like sequences in Solibacter and Zobellia, far better than Nematostella or sponge matches.*	No non-deuterostome metazoan sequences in gene phylogeny (S10)	type IV	removal of sialic acid from novel glycosidic linkages of deuterostomes
13	Proprotein convertase subtilisin/kexin type 9-related, pcsk9 (expanded family in amphioxus and ambulacraria)	*e-59 match to Platynereis (but no inhibitor domain) and e-40 to Capitella (different subtilisin type), and no intron boundary matches*	*e-103 to e-90 for Kexin-like sequences of Actinobacteria such as Saccharothrix and Saccharomonospora*	No non-deuterostome metazoan sequences in gene phylogeny	HGT/convergence	proteolysis, targets unknown (some signaling receptors in vertebrates)
14	PAD (Peptidyl Arginyl Deiminase)	*e+1.7 to an unrelated Ceratitis sequence that lacked PAD domains*	*e-66 match to Stanieria (cyanobacterium) across all domains*	No non-deuterostome metazoan sequences in gene phylogeny	HGT/convergence	post-translational modification of proteins (change arginyl to citrullinyl residues)
15	FATSO (Alpha-ketoglutarate-dependent dioxygenase FTO)		*e-67 to e-60 for matches in Micromonas and Ostreococcus (green microalgae)*	No non-deuterostome metazoan sequences in gene phylogeny (S10)	HGT/convergence	N6-methyladenosine demethylation of nuclear RNAs
16	Arylsulfatase K-like	*e-30 to e-25 matches to Lottia and Capitella (protostomes), but these match much better to a different sulfatase (2-O-sulfo-iduronate sulfatase) with bilaterian ancestry*	*e-83 to e-63 for matches of Monosiga (choanoflagellate) and Arthrobacter (bacterium)*	No non-deuterostome metazoan sequences in gene phylogeny, except distant unrelated sulfatases (S10)	HGT/convergence	removal of sulfate groups, targets unknown
17	polyketide synthase-like	Metazoan fatty acid synthases e-119 to e-94, but different intron-exon structure and domain subtypes.	*e=0.0 (i.e. <e-200) matches to the bacteria Flammeovirga (same order of all domains as deuterostome Family 1), and Sorangium, and the cyanobacterium Stanieria*	No non-deuterostome metazoan sequences in gene phylogeny	HGT/convergence	Synthesis of pigments and antimicrobial agents?
18	NHL-containing protein	e-15 matches to protostomian peptidylglycine alpha-amidating monooxygenases which contains different NHL repeats and a copper monoxygenase domain	*e-38 matches in planctomycete bacteria such as Zavarzinella; and in Monosiga (choanoflagellate)*	No non-deuterostome metazoan sequences in gene phylogeny	HGT/convergence	protein-protein interactions?
19	choline monoxygenase-like	>e-1 unrelated matches	e-97 match to Selaginella (spikemoss), e-91 to Physcomitrella (bryophyte), e-90 to Micromonas, Nannochloropsis, and Coccomyxa (micro-algae)	No non-deuterostome metazoan sequences in gene phylogeny (S10)	HGT/convergence	first step in the synthesis of betaine, an osmotic protectant and also a methyl donor for methionine regeneration
20	ectoine synthase-like (expanded family in amphioxus and hemichordates)	e+7 match unrelated phospholipaseA2 sequence	*e-21 to Rhizobium (bacterium) and Cyanothece (cyanobacterium)*	No non-deuterostome metazoan sequences in gene phylogeny	HGT/convergence	a reaction step in the synthesis of ectoine, an osmotic protectant
21	ectoine hydrolase	*e-18 to e-4 matches to phantoyl dioxygenases of protostomes and deuterosotmes, containing a distantly related domain*	*e-69 to e-45 matches to ectoine hydroxylases of high GC grojup bacteria such as Streptomyces sp. PBH53*	No non-deuterosotme sequences in gene phyogeny	HGT/convergence	a reaction step in the synthesis of betaine, an osmotic protectant
22	histidine methyltransferase bacterial-like (expanded family in amphioxus and hemichordates	*e-34 to e-7 matches to Crassostrea (oyster) sequences, which share 2-3 intron boundaries with deuterostome sequences. No other matches outside deuterostomes*	*e-46 to e-40 matches to bacteria such as Fimbriimonas, Kuenenia, and Thiothrix.*	*Crassostrea sequences more closely associated with deuterostome sequences on gene phylogeny, than are bacterial sequences (S10)*	type IV	Syntheis of trimethyl-histidine derivatives (e.g., ergothionine)?
23	Aromatic amino acid decarboxylase microbial-like (expanded family in amphioxus and hemichordates)	*e-46 to e-10 matches to a different kind of aromatic amino acid decarboxylase in protostomes and deuterostomes, of bilaterian ancestry. Different intron-exon structure.*	*e-157 to e-125 matches to Maribacter (marine bacterium) and other bacteria*	No non-deuterostome metazoan sequences in gene phylogeny (S10)	HGT/convergence	Decarboxylation of aromatic amino acids; neurotransmitter production?
24	Cobalamine-independent methionine synthase	*e-115 match to Nematostella (cnidarian) sequence that shares an intron boundary with the deuterostome sequences, indicating bilaterian ancestry. No protostome match*	*e-119 to e-99 matches to bacteria such as Planctomycetales and Rhodovibrio*	*Nematostella sequence more closely associated with deuterostome sequences on gene phylogeny, than are bacterial sequences*	type IV	Addtional pathway of methonine regeneration in deuterostomes?
25	Major Facilitator Transporter algal-like, MFS algal-like	*e+1.6 match to unrelated cubilin-like protein of Microplitis (insect) with no MFS domain*	*e=0.0 (i.e., <e-200) matches to Chlamydimonas, Ostreococcus, Micromonas, and many other micro-algae*	No non-deuterostome metazoan sequences in gene phylogeny	HGT/convergence	Transmembrane export of noxious compounds?
26	Multicopper oxidase (MCO); also called Bilirubin oxidase-like (expanded family in tunicates, amphioxus, and hemichordates)	*e-5 matches to insect laccases which contain three insect cupredoxin domains that differ greatly from bilirubin oxidase domains*	*e-75 match to Emeliania (coccolithophore) and e-65 to Albugo (oomycete)*	No non-deuterostome metazoan sequences in gene phylogeny	HGT/convergence	oxidation of tetrapyrroles?
27	FAM198-like protein	none	-	Only deuterostome sequenes in gene phylogeny	type I	unknown
28	Chromosome 9 Open Reading Frame 9, Rsb66	*e-15 matches in Xenoturbella, Presagittifera (acoel), Meara (nemertodermatid)*	-	Only deuterostome and acoel/nemertodermatid sequences in gene phylogeny	Tenetatively type I	unknown, expressed in ciliated cells
29	Melanoregulin-like	none	-	Only deuterostome sequenes in gene phylogeny	type I	unknown
30	Small integral membrane protein 19-like	none	-	Only deuterostome sequenes in gene phylogeny	type I	unknown
31	EF-hand and coiled-coil domain-containing protein 1-like	none	-	Only deuterostome sequenes in gene phylogeny	type I	unknown

This table summarizes the description of the novelties in Supplementary Note 10.

ARTICLE

doi:10.1038/nature15530

A perisinusoidal niche for extramedullary haematopoiesis in the spleen

Christopher N. Inra[1]*, Bo O. Zhou[1]*, Melih Acar[1], Malea M. Murphy[1], James Richardson[2], Zhiyu Zhao[1] & Sean J. Morrison[1,3]

Haematopoietic stresses mobilize haematopoietic stem cells (HSCs) from the bone marrow to the spleen and induce extramedullary haematopoiesis (EMH). However, the cellular nature of the EMH niche is unknown. Here we assessed the sources of the key niche factors, SCF (also known as KITL) and CXCL12, in the mouse spleen after EMH induction by myeloablation, blood loss, or pregnancy. In each case, *Scf* was expressed by endothelial cells and *Tcf21*$^+$ stromal cells, primarily around sinusoids in the red pulp, while *Cxcl12* was expressed by a subset of *Tcf21*$^+$ stromal cells. EMH induction markedly expanded the *Scf*-expressing endothelial cells and stromal cells by inducing proliferation. Most splenic HSCs were adjacent to *Tcf21*$^+$ stromal cells in red pulp. Conditional deletion of *Scf* from spleen endothelial cells, or of *Scf* or *Cxcl12* from *Tcf21*$^+$ stromal cells, severely reduced spleen EMH and reduced blood cell counts without affecting bone marrow haematopoiesis. Endothelial cells and *Tcf21*$^+$ stromal cells thus create a perisinusoidal EMH niche in the spleen, which is necessary for the physiological response to diverse haematopoietic stresses.

The haematopoietic system employs facultative niches that arise in response to injury. Adult haematopoiesis occurs primarily in the bone marrow of mammals. However, a wide range of haematopoietic stresses including myelofibrosis[1], anaemia[2,3], pregnancy[4,5], infection[6,7], myeloablation[8] and myocardial infarction[9] can induce EMH, in which HSCs are mobilized to sites outside the bone marrow to expand haematopoiesis. The splenic red pulp is a prominent site of EMH in mice and humans[10–13]. During EMH, HSCs are found mainly around sinusoids in the red pulp, raising the possibility of a perisinusoidal niche[14]. CXCL12 is expressed by sinusoidal endothelial cells in the red pulp of the human spleen[15] and macrophage ablation reduces splenic erythropoiesis after irradiation[16]. However, little else is known about the EMH niche.

Niche factor expression in the spleen

HSCs are rare in normal adult spleen[17] but myeloablation with cyclophosphamide followed by daily administration of granulocyte colony-factor (G-CSF) induces HSC mobilization from the bone marrow to the spleen and induction of EMH[8]. Cyclophosphamide plus 21 days of G-CSF (Cy+21 d G-CSF) increased erythropoiesis and myelopoiesis in the red pulp, profoundly increasing spleen size, spleen cellularity, HSC number and progenitor numbers relative to control spleens (Extended Data Fig. 1c, f–m).

In normal adult spleens from *Scf*^{GFP}; *Cxcl12*^{DsRed} mice[18,19], and after EMH induction, *Scf*-green fluorescent protein (GFP) and *Cxcl12*-DsRed were primarily expressed throughout the red pulp (Fig. 1a, b and Extended Data Fig. 1a–e). Red pulp endothelial cells and perivascular stromal cells expressed high levels of *Scf*-GFP, irrespective of EMH induction (Fig. 1a–c and Extended Data Fig. 1d, e). In white pulp, *Scf*-GFP was expressed by many fewer stromal cells and central arteriolar endothelial cells (Fig. 1b and Extended Data Fig. 1e). *Cxcl12*-DsRed was not expressed by endothelial cells but was expressed by a subset of *Scf*-GFP$^+$ perivascular stromal cells, primarily around red pulp sinusoids and to a lesser extent around white pulp central arterioles (Fig. 1a–c and Extended Data Fig. 1d, e).

Scf-GFP$^+$ cells were $0.48 \pm 0.10\%$ of enzymatically dissociated adult spleen cells (Fig. 1d) and *Cxcl12*-DsRed$^+$ cells were $0.031 \pm 0.011\%$ (Fig. 1f). Most *Scf*-GFP$^+$ cells ($75 \pm 5.8\%$) were VE-cadherin$^+$CD45$^-$Ter119$^-$ endothelial cells (Fig. 1d): $85 \pm 8.2\%$ of all VE-cadherin$^+$CD45$^-$Ter119$^-$ spleen endothelial cells were *Scf*-GFP$^+$ and none expressed *Cxcl12*-DsRed (Fig. 1e). Non-endothelial *Scf*-GFP$^+$ cells were virtually all PDGFR-β$^+$CD45$^-$Ter119$^-$ stromal cells (Fig. 1d). Some *Scf*-GFP$^+$ stromal cells ($22 \pm 3.8\%$) also expressed *Cxcl12*-DsRed (Fig. 1d). Virtually all *Cxcl12*-DsRed$^+$ stromal cells expressed *Scf*-GFP (Fig. 1f). Therefore, *Scf* was expressed by VE-cadherin$^+$ endothelial cells and PDGFR-β$^+$ stromal cells while *Cxcl12* was expressed by a minority of *Scf*-expressing stromal cells in adult spleen.

EMH induction did not appear to alter spleen *Scf*-GFP or *Cxcl12*-DsRed expression (Fig. 1a versus Extended Data Fig. 1d). Flow cytometric analysis showed no change in the fluorescence intensity of individual *Scf*-GFP$^+$ or *Cxcl12*-DsRed$^+$ spleen cells after EMH induction (Extended Data Fig. 1o, p). However, the frequencies and absolute numbers of *Scf*-GFP$^+$ and *Cxcl12*-DsRed$^+$ cells increased significantly upon EMH induction (Fig. 1g–j and Extended Data Fig. 1q, r). These cells rarely divided in normal adult spleen but proliferated upon EMH induction (Fig. 1j, k).

LepR$^+$ stromal cells are the main sources of SCF and CXCL12 for HSC maintenance in the bone marrow[18–20]. In the spleens of *Lepr*^{cre}; *R26*^{tdTomato} mice, recombination occurred mainly in the white pulp, where HSCs are not observed[14] (Extended Data Fig. 1s). Only about 20% of *Scf*-GFP$^+$ stromal cells expressed LepR (Extended Data Fig. 1t). LepR$^+$ cells were PDGFR-β$^+$VE-cadherin$^-$ stromal cells that accounted for $37 \pm 13\%$ of colony-forming-unit fibroblasts (CFU-Fs) formed by enzymatically dissociated spleen cells (Extended Data Fig. 1u, v).

Consistent with our prior study[19], *Lepr*^{cre}; *Scf*$^{fl/-}$ mice had significantly fewer CD150$^+$CD48$^-$LSK HSCs in the bone marrow and significantly increased spleen cellularity relative to *Scf* $^{+/-}$ and *Scf* $^{+/+}$ controls (Extended Data Fig. 1w, x). Upon EMH induction by cyclophosphamide plus 4 days of G-CSF (Cy+4 d G-CSF), *Lepr*^{cre}; *Scf*$^{fl/-}$ mice exhibited significant declines in spleen cellularity and spleen HSC number relative to controls (Extended Data Fig. 1x, y). While LepR$^+$ perivascular stromal cells could contribute to the EMH niche in adult spleen, the impaired EMH in these mice may also reflect bone marrow HSC depletion before EMH induction (Extended Data Fig. 1w).

[1]Department of Pediatrics and Children's Research Institute, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA. [2]Department of Pathology, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA. [3]Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA.
*These authors contributed equally to this work.

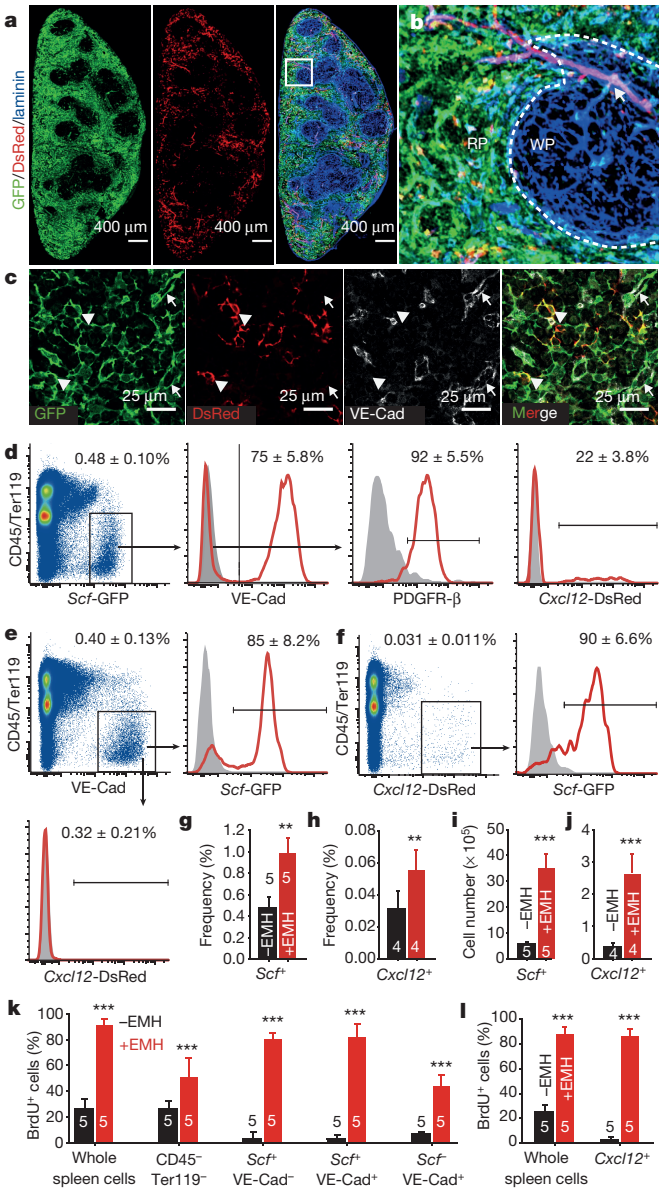


Figure 1 | Endothelial cells and perivascular stromal cells in the red pulp express *Scf* and *Cxcl12* and proliferate upon induction of EMH. **a**, **b**, *Scf*-GFP and *Cxcl12*-DsRed were mainly expressed by stromal cells in the red pulp of normal spleens. **b**, High-magnification view of the boxed area in **a**. Dashed lines depict the boundary between white pulp (WP) and red pulp (RP). Arrow indicates central arteriole in the white pulp, around which rare stromal cells expressed *Cxcl12*-DsRed. **c**, Splenic red pulp from Scf^{GFP}; $Cxcl12^{DsRed}$ mice had VE-cadherin$^+$ endothelial cells (arrows) that expressed *Scf*-GFP and VE-cadherin$^-$ stromal cells (arrowheads) that expressed *Scf*-GFP and sometimes *Cxcl12*-DsRed. VE-Cad, VE-cadherin. **d**–**f**, Flow cytometric analysis of enzymatically dissociated spleen cells from Scf^{GFP}; $Cxcl12^{DsRed}$ mice. *Scf*-GFP was expressed by VE-cadherin$^+$ endothelial cells and PDGFR-β$^+$VE-cadherin$^-$ stromal cells, a subset of which also expressed *Cxcl12*-DsRed (**d**). Most VE-cadherin$^+$ endothelial cells were positive for *Scf*-GFP but negative for *Cxcl12*-DsRed (**e**). Most *Cxcl12*-DsRed$^+$ cells were positive for *Scf*-GFP (**f**). Data in **a**–**f** represent mean ± s.d. from 3 mice from 3 independent experiments. **g**–**l**, The frequencies and absolute numbers of *Scf*-GFP$^+$ cells (**g**, **i**) and *Cxcl12*-DsRed$^+$ cells (**h**, **j**) significantly increased upon induction of EMH by Cy+21 d G-CSF (+EMH). **k**, **l**, 5-Bromo-2′-deoxyuridine (BrdU) was co-administered to Scf^{GFP} (**k**) or $Cxcl12^{DsRed}$ mice (**l**) along with G-CSF for 7 days after cyclophosphamide treatment. Data represent mean ± s.d. from 3 independent experiments. The numbers of mice per treatment are shown on the bars in panels. **g**–**l**, Two-tailed Student's *t*-tests were used to assess statistical significance (**$P < 0.01$, ***$P < 0.001$).

Tcf21$^+$ perisinusoidal stromal cells express *Scf*

To identify *cre* alleles that recombine in spleen, but not bone marrow, stromal cells, we assessed the gene expression profile of spleen *Scf*-GFP$^+$VE-cadherin$^-$ stromal cells (Extended Data Table 1). After testing a number of *cre* alleles (see Extended Data Fig. 2), we found that *Tcf21*-Cre/ER (ref. 21) recombined efficiently in spleen *Scf*-GFP$^+$ stromal cells (Fig. 2a) but not in bone marrow (Fig. 2b, c). $Tcf21^{cre/ER}$; $R26^{tdTomato}$ mice gavaged with tamoxifen for 12 days at 4–6 weeks of age expressed Tomato in *Scf*-GFP$^+$ stromal cells throughout the red pulp (Fig. 2a, d), whereas Tomato was expressed only in rare white pulp cells (Fig. 2a) and in no endothelial cells (Fig. 2d, e). Tomato$^+$CD45$^-$Ter119$^-$ stromal cells from enzymatically dissociated $Tcf21^{cre/ER}$; $R26^{tdTomato}$ spleens accounted for 0.085 ± 0.045% of spleen cells and 69 ± 2% of spleen CFU-Fs (Fig. 2f, g). These cells were PDGFR-β$^+$ and LepR$^-$ (Fig. 2f).

In the liver, *Scf*-GFP was exclusively expressed by VE-cadherin$^+$ endothelial cells (Extended Data Fig. 2a, b). *Tcf21*-Cre/ER recombined in 0.09% of liver cells, none of which expressed *Scf*-GFP (Extended Data Fig. 2a, c). The *Tcf21*-Cre/ER recombination pattern did not significantly change in the spleen (Fig. 2f and Extended Data Fig. 2d, e), bone marrow (Extended Data Fig. 2f, g), or liver (Extended Data Fig. 2h, i) upon EMH induction by Cy+21 d G-CSF.

c-Kit$^+$ haematopoietic progenitors were almost exclusively within the red pulp in the normal spleen (Extended Data Fig. 3a, b) and after EMH induction (Fig. 2k). To assess HSC localization we used a new technique that permits deep imaging of α-*catulin*-GFP$^+$c-Kit$^+$ HSCs in optically cleared haematopoietic tissues[22]. In the spleens of mice treated with Cy+4 d G-CSF, only 0.019 ± 0.01% of splenocytes were α-*catulin*-GFP$^+$c-Kit$^+$ (Fig. 2h). All long-term multilineage reconstituting cells in the spleen were α-*catulin*-GFP$^+$ and 28% of α-*catulin*-GFP$^+$c-Kit$^+$ spleen cells gave long-term multilineage reconstitution in primary (Fig. 2i) and secondary irradiated recipient mice (data not shown).

After antibody staining of a large segment of $Tcf21^{cre/ER}$; $R26^{tdTomato}$; $\alpha\text{-}catulin^{GFP}$ spleen, we cleared the tissue (Extended Data Fig. 3c, d), then imaged to a depth of 300 μm and digitally reconstructed the tissue (Extended Data Fig. 3e, f and Supplementary Video 1). α-*Catulin*-GFP$^+$c-Kit$^+$ HSCs were found exclusively within the red pulp, where 80% were within 5 μm of Tomato$^+$ stromal cells (Fig. 2j).

EMH requires SCF and CXCL12 from *Tcf21*$^+$ cells

To test whether $Tcf21^{cre/ER}$-expressing perivascular cells promote EMH, we treated 4–6-week-old $Tcf21^{cre/ER}$; $Scf^{fl/fl}$ and littermate control mice with tamoxifen for 12 days. A month later, bone marrow and spleen cellularity, blood cell counts, and bone marrow haematopoiesis were similar in $Tcf21^{cre/ER}$; $Scf^{fl/fl}$ mice and littermate controls (Fig. 3a–f and Extended Data Fig. 3g–l). Then we treated $Tcf21^{cre/ER}$; $Scf^{fl/fl}$ mice and littermate controls with cyclophosphamide followed by 4, 8, or 21 days of G-CSF. $Tcf21^{cre/ER}$; $Scf^{fl/fl}$ mice did not differ from controls with respect to bone marrow cellularity (Fig. 3a) or the numbers of HSCs (Fig. 3b), common myeloid progenitors (CMPs[23]), granulocyte–macrophage progenitors (GMPs[23]), or megakaryocyte–erythroid progenitors (MEPs[23]) in the bone marrow after Cy+4–21 d G-CSF treatment (Extended Data Fig. 3j–l). In contrast, $Tcf21^{cre/ER}$; $Scf^{fl/fl}$ mice had significantly fewer splenocytes (Fig. 3c), spleen HSCs (Fig. 3d), CMPs (Fig. 3e), GMPs (Extended Data Fig. 3m) and MEPs (Fig. 3f) relative to littermate controls after Cy+8–21 d G-CSF treatment. We did not detect any difference between $Tcf21^{cre/ER}$; $Scf^{fl/fl}$ mice and littermate controls in terms of vascular or stromal cell morphology in the spleen, with or without induction of EMH (Extended Data Fig. 4a–g). Conditional deletion of *Scf* with *Tcf21*-Cre/ER thus depletes HSCs and reduces EMH in the spleen without affecting bone marrow haematopoiesis.

Red blood cell (RBC) and white blood cell (WBC) counts were significantly lower in $Tcf21^{cre/ER}$; $Scf^{fl/fl}$ mice as compared to controls after Cy+8–21 d G-CSF treatment (Extended Data Fig. 3g–i). Splenectomy significantly reduced RBC and WBC counts in mice treated with Cy+G-CSF, demonstrating that splenic EMH is necessary for the

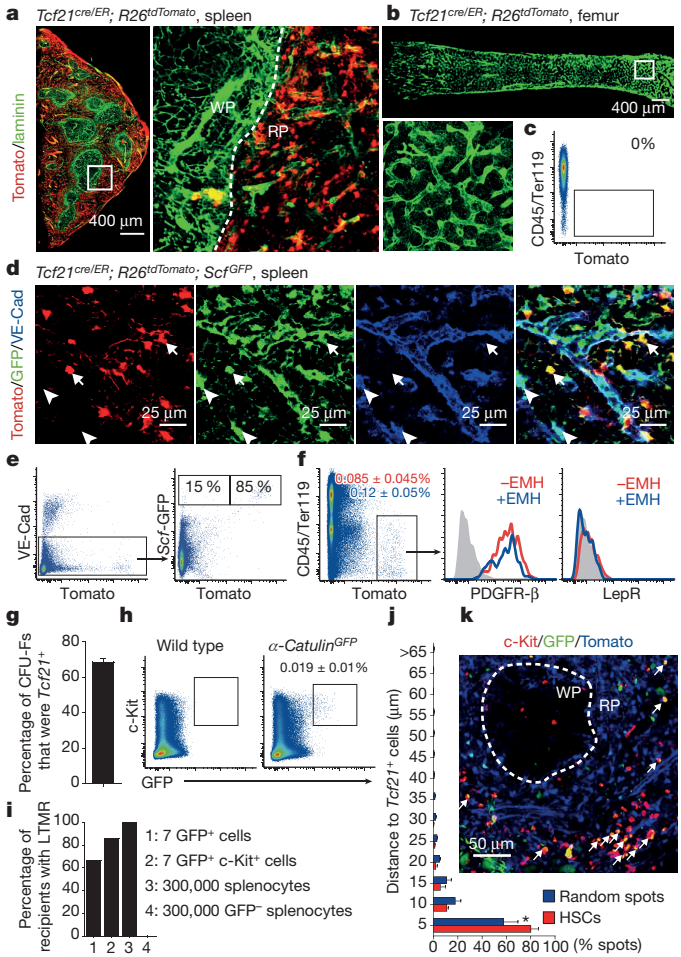


Figure 2 | During EMH most HSCs localize adjacent to *Tcf21*$^+$ stromal cells in the red pulp. a, Tamoxifen-treated adult *Tcf21*$^{cre/ER}$; *R26*^{tdTomato} mice exhibited widespread Tomato expression by perivascular stromal cells in the red pulp (RP). **b**, **c**, No Tomato expression in bone marrow from tamoxifen-treated *Tcf21*$^{cre/ER}$; *R26*^{tdTomato} mice. **d**, **e**, Most *Scf*-GFP$^+$VE-cadherin$^-$ stromal cells were Tomato$^+$ (arrows) whereas *Scf*-GFP$^+$VE-cadherin$^+$ endothelial cells were Tomato$^-$ (arrowhead). **f**, Tomato$^+$CD45$^-$Ter119$^-$ stromal cells from enzymatically dissociated spleen from *Tcf21*$^{cre/ER}$; *R26*^{tdTomato} mice were positive for PDGFR-β but negative for LepR, irrespective of EMH induction by Cy+G-CSF. **g**, Percentage of all CFU-F colonies formed by enzymatically dissociated *Tcf21*$^{cre/ER}$; *R26*^{tdTomato} spleen cells that were Tomato$^+$. Macrophage colonies were excluded by staining with anti-CD45 antibody. **h**, α-*Catulin*-GFP$^+$c-Kit$^+$ HSCs represented 0.019 ± 0.01% of dissociated spleen cells in α-*catulin*^{GFP} mice with EMH. **i**, α-*Catulin*-GFP$^+$c-Kit$^+$ splenocytes were highly enriched for long-term multilineage reconstituting (LTMR) HSCs. **j**, **k**, Deep imaging of α-*catulin*–GFP$^+$c-Kit$^+$ HSCs (arrows in **k**) in optically cleared spleen from a *Tcf21*$^{cre/ER}$; *R26*^{tdTomato}; α-*catulin*^{GFP} mouse with EMH induced by Cy+21 d G-CSF. The distance from α-*catulin*-GFP$^+$c-Kit$^+$ HSCs or random spots to Tomato$^+$ stromal cells (**j**; $^*P < 0.05$ by two-tailed Student's *t*-test). α-*Catulin*-GFP$^+$c-Kit$^+$ HSCs were exclusively in the red pulp (**k**; see Extended Data Fig. 3f for a low-magnification view). All data reflect mean ± s.d. from 3 mice in 3 independent experiments.

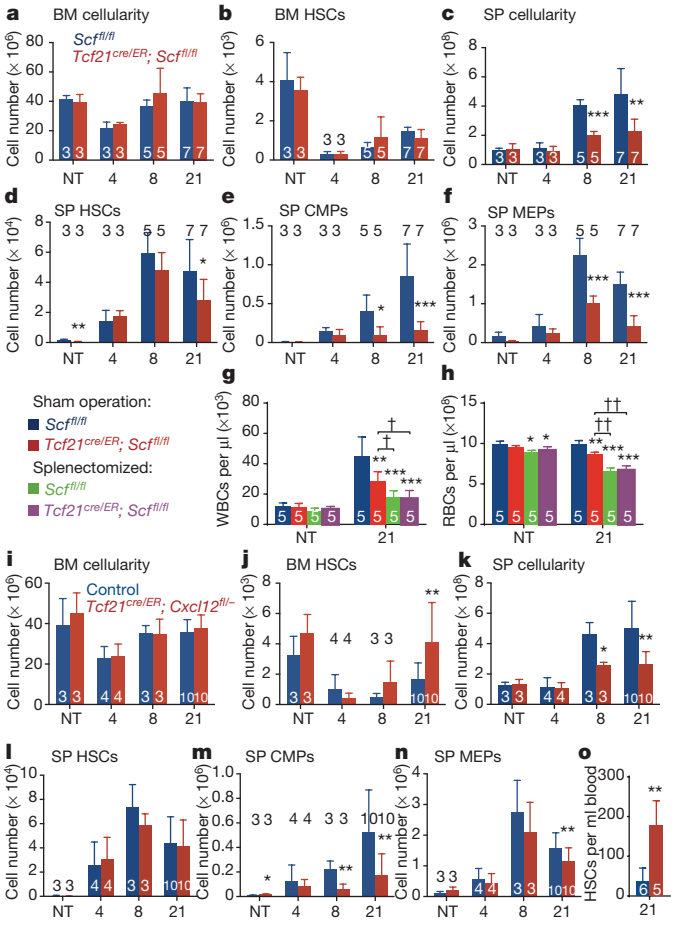


Figure 3 | *Tcf21*-expressing stromal cells are an important source of SCF and CXCL12 for EMH in the spleen. a–**f**, *Tcf21*$^{cre/ER}$; *Scf*$^{fl/fl}$ and *Scf*$^{fl/fl}$ control mice were treated with tamoxifen then examined 1 month later either under normal conditions (not treated (NT)) or after treatment with Cy+4–21 d G-CSF to induce EMH. The number of bone marrow (BM) cells (**a**) and bone marrow CD150$^+$CD48$^-$LSK HSCs (**b**) in one femur plus one tibia as well as spleen (SP) cellularity (**c**) and the numbers of HSCs (**d**), CMPs (**e**) and MEPs (**f**) in the spleen. **g**, **h**, Sham-operated and splenectomized mice were treated with Cy+21d G-CSF 1 month after surgery: WBC (**g**) and RBC (**h**) counts are shown. **i**–**o**, *Tcf21*$^{cre/ER}$; *Cxcl12*$^{fl/-}$ and *Cxcl12*$^{+/-}$ or *Cxcl12*$^{fl/-}$ control mice were treated with tamoxifen then examined 1 month later either under normal conditions (NT) or after treatment with Cy+4–21d G-CSF to induce EMH. The number of bone marrow cells (**i**) and bone marrow HSCs (**j**) in one femur plus one tibia as well as spleen cellularity (**k**), numbers of HSCs (**l**), CMPs (**m**) and MEPs (**n**) in the spleen are shown. **o**, Number of HSCs per ml of blood in tamoxifen-treated control and *Tcf21*$^{cre/ER}$; *Cxcl12*$^{fl/-}$ mice after Cy+21 d G-CSF. The numbers of mice per treatment are shown in each bar in each panel. All panels reflect mean ± s.d. from three independent experiments. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, statistical significance relative to sham-operated *Scf*$^{fl/fl}$ mice. †$P < 0.05$, ††$P < 0.01$, statistical significance among other treatments.

recovery of blood cell counts (Fig. 3g, h and Extended Data Fig. 3n). However, conditional deletion of *Scf* by *Tcf21*-Cre/ER did not further reduce blood cell counts in splenectomized mice (Fig. 3g, h). SCF expression by *Tcf21*$^+$ stromal cells in the spleen is thus necessary for the regeneration of blood cells after Cy+G-CSF treatment.

Bone marrow cellularity and bone marrow haematopoiesis were similar in *Tcf21*$^{cre/ER}$; *Cxcl12*$^{fl/-}$ mice and littermate controls, before and after Cy+G-CSF treatment (Fig. 3i, j and Extended Data Fig. 3r–t). However, *Tcf21*$^{cre/ER}$; *Cxcl12*$^{fl/-}$ mice exhibited significantly reduced spleen cellularity (Fig. 3k) and numbers of spleen CMPs, GMPs and MEPs (Fig. 3m, n and Extended Data Fig. 3u) relative to controls after Cy+8–21 d G-CSF treatment. Although the number of HSCs in the spleens of *Tcf21*$^{cre/ER}$; *Cxcl12*$^{fl/-}$ mice did not significantly differ from littermate controls (Fig. 3l), HSC numbers were significantly elevated in the blood (Fig. 3o) and in the bone marrow (Fig. 3j) of *Tcf21*$^{cre/ER}$; *Cxcl12*$^{fl/-}$ mice after Cy+21 d G-CSF treatment. This suggests that some HSCs were mobilized from the spleens of *Tcf21*$^{cre/ER}$; *Cxcl12*$^{fl/-}$ mice. *Tcf21*$^{cre/ER}$; *Cxcl12*$^{fl/-}$ mice also had significantly reduced RBC counts after Cy+21 d G-CSF treatment (Extended Data Fig. 3o–q). We did not detect any difference between *Tcf21*$^{cre/ER}$; *Cxcl12*$^{fl/-}$ mice and littermate controls in terms of the frequency or morphology of vascular or stromal cells in the spleen, with or without EMH

(Extended Data Fig. 4h–n). *Tcf21*-Cre/ER-expressing stromal cells are thus an important source of CXCL12 for spleen EMH but not bone marrow haematopoiesis.

EMH requires SCF from endothelial cells

We discovered that *Vav1*-Cre recombines efficiently in spleen, but not bone marrow, endothelial cells. *Vav1-cre; R26*^{tdTomato} mice recombined throughout the red pulp in VE-cadherin^{+}*Scf*-GFP^{+} cells but only in rare white pulp cells (Fig. 4a–c). VE-cadherin^{+}*Scf*-GFP^{+} cells accounted for 0.37 ± 0.07% of enzymatically dissociated spleen cells and 83 ± 5.3% of these cells recombined with *Vav1*-Cre (Fig. 4b). These cells were negative for PDGFR-β (Extended Data Fig. 5a). Seventy ± 5% of VE-cadherin^{+} endothelial cells were Tomato^{+} in the spleens of *Vav1-cre; R26*^{tdTomato} mice but only 8.4 ± 0.5% were Tomato^{+} in bone marrow (Extended Data Fig. 5b, e–h). Endothelial cells from *Vav1-cre; Scf*$^{fl/-}$ mice exhibited a 6.5-fold reduction in *Scf* transcript levels (Extended Data Fig. 5c) and a 5.6-fold reduction in SCF protein (Extended Data Fig. 5d) relative to endothelial cells from *Scf*$^{fl/-}$ controls.

In the livers of *Vav1-cre; R26*^{tdTomato}; *Scf*^{GFP} mice recombination occurred in 26 ± 4.2% of VE-cadherin^{+}*Scf*-GFP^{+} cells (Extended Data Fig. 5i–k). Upon induction of EMH by Cy+G-CSF, *Vav1*-Cre recombination did not significantly change in the spleen (Extended Data Figs 5b and 6a, b), bone marrow (Extended Data Figs 5b and 6c, d) or liver (Extended Data Fig. 6e, f).

Cxcl12 was not expressed by spleen endothelial cells (Fig. 1e). Consistent with this, *Vav1-cre; Cxcl12*$^{fl/-}$ mice had normal blood counts, cellularity, and numbers of HSCs, CMPs, GMPs and MEPs in bone marrow and spleen after Cy+G-CSF (Extended Data Fig. 6g–s).

Vav1-Cre also recombines in haematopoietic cells[24] but haematopoietic cells do not express *Scf* and *Vav1-cre; Scf*$^{fl/-}$ mice have normal HSC frequency and haematopoiesis in bone marrow[18,19]. Prior to EMH induction with Cy+G-CSF, *Vav1-cre; Scf*$^{fl/-}$ mice did not significantly differ from *Scf*$^{fl/-}$ controls with respect to bone marrow or spleen cellularity, or the numbers of HSCs, CMPs, GMPs or MEPs in the bone marrow or spleen (Fig. 4d–i and Extended Data Fig. 6w–z). After Cy+G-CSF treatment, bone marrow cellularity and numbers of bone marrow HSCs, CMPs, GMPs or MEPs in *Vav1-cre; Scf*$^{fl/-}$ mice were normal (Fig. 4d, e and Extended Data Fig. 6w–y). However, RBC counts, spleen cellularity, and the numbers of spleen HSCs, CMPs and MEPs declined in *Vav1-cre; Scf*$^{fl/-}$ mice relative to *Scf*$^{fl/-}$ controls (Fig. 4f–i and Extended Data Fig. 6t–v).

The decline in blood cell counts in *Vav1-cre; Scf*$^{fl/-}$ mice after EMH induction was caused by reduced spleen EMH because splenectomy significantly reduced RBC and WBC counts but conditional deletion of *Scf* in splenectomized *Vav1-cre; Scf*$^{fl/-}$ mice had no further effect on blood cell counts (Fig. 4j, k). We did not detect any difference between *Vav1-cre; Scf*$^{fl/-}$ mice and controls in terms of the frequency or morphology of vascular or stromal cells in the spleen (Extended Data Fig. 4o–u). Endothelial SCF expression is thus necessary for splenic EMH and the recovery of blood cell counts after Cy+G-CSF.

The splenic EMH niche during pregnancy

Erythropoiesis and myelopoiesis significantly increased in the red pulp during pregnancy, profoundly increasing spleen cellularity, HSC number, and progenitor numbers relative to non-pregnant mice (Extended Data Fig. 7a–i). Just as in Cy+G-CSF-treated mice, *Scf*-GFP was largely expressed by endothelial and perivascular stromal cells in the red pulp and *Cxcl12*-DsRed was expressed by a subset of the *Scf*-GFP^{+} stromal cells (Extended Data Fig. 7j–l). Pregnancy induced these cells to proliferate, significantly expanding their numbers (Extended Data Fig. 7m–o). In pregnant mice, *Tcf21*-Cre/ER recombined in spleen PDGFR-β^{+}LepR^{-} stromal cells but not in bone marrow and rarely in liver (Extended Data Fig. 7p–v). *Vav1-cre; Scf*$^{fl/fl}$ mice were infertile, preventing us from testing the endothelial contribution to EMH during pregnancy.

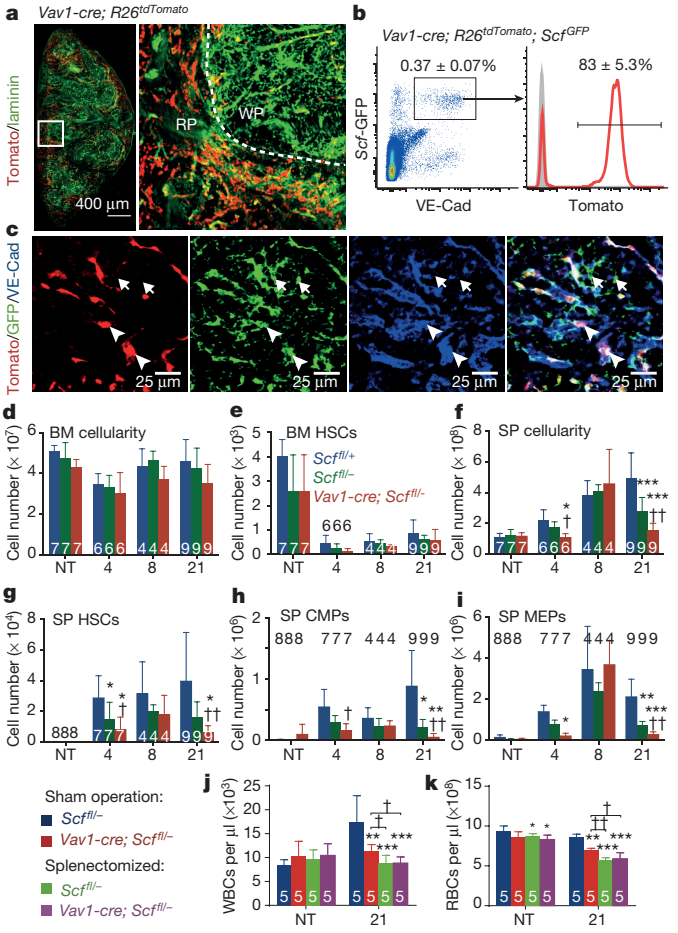


Figure 4 | Endothelial cells are an important source of SCF for EMH in the spleen. a, *Vav1-cre; R26*^{tdTomato} mice exhibited vascular Tomato expression throughout the splenic red pulp (RP). Tomato was also expressed by haematopoietic cells in these mice but levels of Tomato expression in endothelial cells were ~10–100 fold higher than in haematopoietic cells. Therefore short-exposure images showed mainly Tomato fluorescence in endothelial cells. WP, white pulp. **b**, **c**, *Vav1*-Cre recombined in VE-cadherin^{+}*Scf*-GFP^{+} endothelial cells (arrowheads in **c**) but not in VE-cadherin^{-}*Scf*-GFP^{+} perivascular stromal cells (arrows in **c**). **d**–**i**, *Vav1-cre; Scf*$^{fl/-}$ mice and *Scf*$^{fl/+}$, *Scf*$^{fl/-}$ controls were not treated (NT) or treated with Cy+4–21 d G-CSF to induce EMH. Data show the number of bone marrow (BM) cells (**d**) and bone marrow HSCs (**e**) in one femur plus one tibia as well as spleen (SP) cellularity (**f**) and the numbers of HSCs (**g**), CMPs (**h**) and MEPs (**i**) in the spleen. **j**, **k**, WBC (**j**) and RBC (**k**) counts in splenectomized and sham-operated mice before and after Cy+21 d G-CSF treatment. The numbers of mice per treatment are shown in the bars in each panel. All data reflect mean ± s.d. from 3 (**a**–**c**; **j**, **k**) or 6 (**d**–**i**) independent experiments. **d**–**i**, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, statistical significance relative to *Scf*$^{fl/+}$. †$P < 0.05$, ††$P < 0.01$, †††$P < 0.001$, statistical significance between *Scf*$^{fl/-}$ and *Vav1-cre; Scf*$^{fl/-}$ mice. **j**, **k**, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, statistical significance relative to sham-operated *Scf*$^{fl/-}$. †$P < 0.05$, ††$P < 0.01$, statistical significance between other treatments.

Pregnant *Tcf21*$^{cre/ER}$; *Scf*$^{fl/fl}$ females did not differ from *Scf*$^{fl/fl}$ control females in terms of bone marrow cellularity (Fig. 5a), or the numbers of HSCs (Fig. 5b), GMPs, CMPs or MEPs in the bone marrow (Extended Data Fig. 8a–d). In contrast, pregnant *Tcf21*$^{cre/ER}$; *Scf*$^{fl/fl}$ females exhibited significantly lower spleen cellularity and numbers of HSCs, GMPs, CMPs, MEPs, myeloid and erythroid cells in the spleen as compared to pregnant *Scf*$^{fl/fl}$ females (Fig. 5c–f and Extended Data Fig. 8e, f). Pregnant *Tcf21*$^{cre/ER}$; *Scf*$^{fl/fl}$ females had significantly lower RBC counts than pregnant *Scf*$^{fl/fl}$ controls (Fig. 5g), and significantly lower fetal mass (Fig. 5h). SCF from *Tcf21*$^{+}$ perivascular cells is thus necessary for splenic EMH and for the expansion of erythropoiesis during pregnancy.

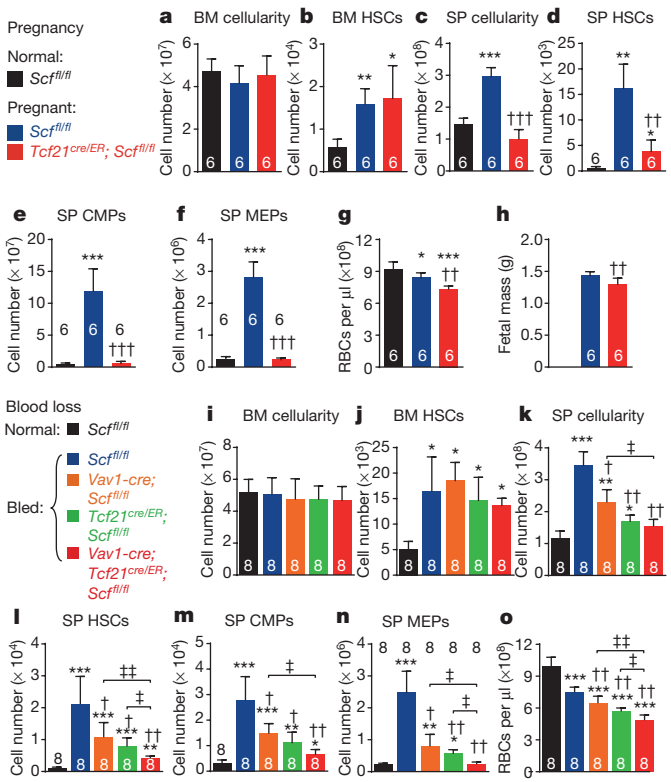


Figure 5 | SCF from endothelial cells and *Tcf21*$^+$ stromal cells is necessary for splenic EMH and adequate erythropoiesis after bleeding or during pregnancy. **a**–**h**, Four-to-six-month-old female mice that had been treated with tamoxifen at least 2 months earlier were mated with normal wild-type males. **a**–**f**, Normal females and pregnant females at gestation day 18.5 were analysed: the number of bone marrow (BM) cells (**a**) and bone marrow HSCs (**b**) in one femur plus one tibia as well as spleen (SP) cellularity (**c**), and the numbers of HSCs (**d**), CMPs (**e**) and MEPs (**f**) in the spleen are shown. **g**, **h**, RBC counts (**g**) and fetal mass (**h**). **i**–**n**, Four-to-six-month-old mice with the indicated genetic backgrounds were repeatedly bled over a 2-week period then analysed: the number of bone marrow cells (**i**) and bone marrow HSCs (**j**) in one femur plus one tibia as well as spleen cellularity (**k**), and the numbers of HSCs (**l**), CMPs (**m**) and MEPs (**n**) in the spleen are shown. **o**, RBC counts. The numbers of mice per treatment are shown in each bar of each panel. All data reflect mean ± s.d. from 4 (**a**–**h**) or 3 (**i**–**o**) independent experiments. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, statistical significance relative to normal mice. $\dagger P < 0.05$, $\dagger\dagger P < 0.01$, statistical significance between single mutants and compound mutants. $\ddagger P < 0.05$, $\ddagger\ddagger P < 0.01$, $\ddagger\ddagger\ddagger P < 0.001$, statistical significance between *Scf* mutant mice and control mice after bleeding or pregnancy.

Pregnant *Tcf21*$^{cre/ER}$; *Cxcl12*$^{fl/-}$ females also had significantly reduced splenic cellularity and splenic erythropoiesis relative to pregnant *Cxcl12*$^{fl/-}$ controls, without any changes in bone marrow haematopoiesis (Extended Data Fig. 8i–x).

The splenic EMH niche after blood loss

Repeated bleeding significantly increased erythropoiesis and myelopoiesis in the red pulp, increasing spleen cellularity, HSC number, and progenitor numbers relative to non-bled controls (Extended Data Fig. 9a–i). Just as in Cy+G-CSF-treated mice, *Scf*-GFP was largely expressed by endothelial cells and perivascular stromal cells in the red pulp while *Cxcl12*-DsRed was expressed by a subset of *Scf*-GFP$^+$ stromal cells (Extended Data Fig. 9j–l). Blood loss induced the proliferation of these cells, significantly expanding their numbers (Extended Data Fig. 9m–o). In bled mice, *Tcf21*-Cre/ER recombined in red pulp PDGFR-β$^+$LepR$^-$ stromal cells, but not in bone marrow and rarely in liver (Extended Data Fig. 9p–v). *Vav1*-Cre recombined in 66 ± 4.2% of spleen endothelial cells, mainly in the red pulp, but only in 7.5 ± 4.0% of bone marrow endothelial cells and 25 ± 5.8% of liver endothelial cells (Extended Data Fig. 10a–h).

Bled *Tcf21*$^{cre/ER}$; *Scf*$^{fl/fl}$ mice or *Vav1-cre; Scf*$^{fl/fl}$ mice did not differ from bled *Scf*$^{fl/fl}$ controls in bone marrow cellularity (Fig. 5i), or the numbers of HSCs (Fig. 5j), GMPs, CMPs or MEPs in the bone marrow (Extended Data Fig. 10i–l). In contrast, bled *Tcf21*$^{cre/ER}$; *Scf*$^{fl/fl}$ mice and *Vav1-cre; Scf*$^{fl/fl}$ mice each had significantly lower RBC counts, spleen cellularity, and numbers of HSCs, GMPs, CMPs, MEPs, myeloid and erythroid cells in the spleen as compared to bled *Scf*$^{fl/fl}$ controls (Fig. 5l–o and Extended Data Fig. 10m–p). *Tcf21*$^+$ stromal cells and endothelial cells are thus necessary for EMH in the spleen and for the expansion of erythropoiesis after bleeding.

Endothelial and *Tcf21*$^+$ stromal cells had additive effects on splenic EMH and the recovery of RBC counts after bleeding. Bled *Vav1-cre; Tcf21*$^{cre/ER}$; *Scf*$^{fl/fl}$ mice had similar bone marrow cellularity and numbers of HSCs in the bone marrow as bled *Scf*$^{fl/fl}$ controls (Fig. 5i, j). However, they had significantly reduced RBC counts, spleen cellularity, and numbers of HSCs, MEPs and erythroid cells in the spleen as compared to bled *Scf*$^{fl/fl}$ mice, bled *Vav1-cre; Scf*$^{fl/fl}$ mice, and bled *Tcf21*$^{cre/ER}$; *Scf*$^{fl/fl}$ mice (Fig. 5k–n and Extended Data Fig. 10p).

Bled *Tcf21*$^{cre/ER}$; *Cxcl12*$^{fl/-}$ mice also had significantly reduced cellularity, MEPs and erythroid cells in the spleen as well as significantly reduced RBC counts as compared to bled *Cxcl12*$^{fl/-}$ controls, without any differences in bone marrow haematopoiesis (Extended Data Fig. 10q–e′).

The EMH niche in mouse spleen is created by endothelial cells and *Tcf21*-expressing stromal cells associated with red pulp sinusoids and is functionally important for haematopoietic recovery from a range of stresses. A prior study[15] detected CXCL12 expression in endothelial cells in human spleens. This suggests that endothelial cells are also a component of the EMH niche in humans, but there may be species differences in CXCL12 expression among niche cells. It is not clear whether there is any relationship between the *Cxcl12*-abundant reticular (CAR) cells that are part of the bone marrow niche[25] and the *Cxcl12*-expressing stromal cells in the splenic EMH niche. While bone marrow CAR cells are LepR$^+$ and *Tcf21*$^-$, spleen CAR cells are *Tcf21*$^+$ and LepR$^-$.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 5 July; accepted 1 September 2015.
Published online 16 November 2015.

1. Abdel-Wahab, O. I. & Levine, R. L. Primary myelofibrosis: update on definition, pathogenesis, and treatment. *Annu. Rev. Med.* **60,** 233–245 (2009).
2. Cheshier, S. H., Prohaska, S. S. & Weissman, I. L. The effect of bleeding on hematopoietic stem cell cycling and self-renewal. *Stem Cells Dev.* **16,** 707–718 (2007).
3. Bennett, M., Pinkerton, P. H., Cudkowicz, G. & Bannerman, R. M. Hemopoietic progenitor cells in marrow and spleen of mice with hereditary iron deficiency anemia. *Blood* **32,** 908–921 (1968).
4. Nakada, D. *et al.* Oestrogen increases haematopoietic stem-cell self-renewal in females and during pregnancy. *Nature* **505,** 555–558 (2014).
5. Fowler, J. H. & Nash, D. J. Erythropoiesis in the spleen and bone marrow of the pregnant mouse. *Dev. Biol.* **18,** 331–353 (1968).
6. Baldridge, M. T., King, K. Y., Boles, N. C., Weksberg, D. C. & Goodell, M. A. Quiescent haematopoietic stem cells are activated by IFN-γ in response to chronic infection. *Nature* **465,** 793–797 (2010).
7. Burberry, A. *et al.* Infection mobilizes hematopoietic stem cells through cooperative NOD-like receptor and Toll-like receptor signaling. *Cell Host Microbe* **15,** 779–791 (2014).
8. Morrison, S. J., Wright, D. E. & Weissman, I. L. Cyclophosphamide/granulocyte colony-stimulating factor induces hematopoietic stem cells to proliferate prior to mobilization. *Proc. Natl Acad. Sci. USA* **94,** 1908–1913 (1997).
9. Dutta, P. *et al.* Myocardial infarction accelerates atherosclerosis. *Nature* **487,** 325–329 (2012).
10. Lowell, C. A., Niwa, M., Soriano, P. & Varmus, H. E. Deficiency of the Hck and Src tyrosine kinases results in extreme levels of extramedullary hematopoiesis. *Blood* **87,** 1780–1792 (1996).
11. Freedman, M. H. & Saunders, E. F. Hematopoiesis in the human spleen. *Am. J. Hematol.* **11,** 271–275 (1981).

12. Tavassoli, M. & Weiss, L. An electron microscopic study of spleen in myelofibrosis with myeloid metaplasia. *Blood* **42,** 267–279 (1973).
13. Johns, J. L. & Christopher, M. M. Extramedullary hematopoiesis: a new look at the underlying stem cell niche, theories of development, and occurrence in animals. *Vet. Pathol.* **49,** 508–523 (2012).
14. Kiel, M. J. *et al.* SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells. *Cell* **121,** 1109–1121 (2005).
15. Miwa, Y. *et al.* Up-regulated expression of CXCL12 in human spleens with extramedullary haematopoiesis. *Pathology* **45,** 408–416 (2013).
16. Chow, A. *et al.* CD169^{+} macrophages provide a niche promoting erythropoiesis under homeostasis and stress. *Nature Med.* **19,** 429–436 (2013).
17. Morita, Y. *et al.* Functional characterization of hematopoietic stem cells in the spleen. *Exp. Hematol.* **39,** 351–359 (2011).
18. Ding, L. & Morrison, S. J. Haematopoietic stem cells and early lymphoid progenitors occupy distinct bone marrow niches. *Nature* **495,** 231–235 (2013).
19. Ding, L., Saunders, T. L., Enikolopov, G. & Morrison, S. J. Endothelial and perivascular cells maintain haematopoietic stem cells. *Nature* **481,** 457–462 (2012).
20. Zhou, B. O., Yue, R., Murphy, M. M., Peyer, J. G. & Morrison, S. J. Leptin-receptor-expressing mesenchymal stromal cells represent the main source of bone formed by adult bone marrow. *Cell Stem Cell* **15,** 154–168 (2014).
21. Acharya, A., Baek, S. T., Banfi, S., Eskiocak, B. & Tallquist, M. D. Efficient inducible Cre-mediated recombination in Tcf21 cell lineages in the heart and kidney. *Genesis* **49,** 870–877 (2011).
22. Acar, M. *et al.* Deep imaging of bone marrow shows non-dividing stem cells are mainly perisinusoidal. *Nature* **526,** 126–130 (2015).
23. Akashi, K., Traver, D., Miyamoto, T. & Weissman, I. L. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature* **404,** 193–197 (2000).
24. de Boer, J. *et al.* Transgenic mice with hematopoietic and lymphoid specific expression of Cre. *Eur. J. Immunol.* **33,** 314–325 (2003).
25. Sugiyama, T., Kohara, H., Noda, M. & Nagasawa, T. Maintenance of the hematopoietic stem cell pool by CXCL12-CXCR4 chemokine signaling in bone marrow stromal cell niches. *Immunity* **25,** 977–988 (2006).

Supplementary Information is available in the online version of the paper.

Acknowledgements S.J.M. is a Howard Hughes Medical Institute Investigator, the Mary McDermott Cook Chair in Pediatric Genetics, the director of the Hamon Laboratory for Stem Cells and Cancer, and a Cancer Prevention and Research Institute of Texas Scholar. B.O.Z. was supported by a fellowship from the Leukemia and Lymphoma Society. We thank N. Loof and the Moody Foundation Flow Cytometry Facility, K. Correll and M. Gross for mouse colony management, and E. Olson and J. Mendell for providing Cre lines. This work was supported by the National Institutes of Health National Heart, Lung, and Blood Institute (HL097760).

Author Contributions C.N.I. identified the *cre* alleles used in this study and analysed *Scf* and *Cxcl12* conditional knockout mice after Cy+G-CSF treatment. B.O.Z. characterized the stromal cells in the spleen and analysed *Scf* and *Cxcl12* conditional knockout mice after blood loss and pregnancy. M.A. generated and characterized the α-*catulin*^{GFP} mice. M.M.M. analysed HSC localization in the spleen. Z.Z. performed all statistical analyses. J.R. examined spleen histology. C.N.I., B.O.Z., M.A., M.M.M. and S.J.M. designed the experiments and interpreted the results. C.N.I., B.O.Z. and S.J.M. wrote the manuscript.

Author Information Microarray data have been deposited in the Gene Expression Omnibus under accession number GSE71288. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.J.M. (sean.morrison@utsouthwestern.edu).

METHODS

Mice. All mice were maintained on a C57BL/6 background, including *Scf*^{GFP} (ref. 19), *Scf*$^{fl/+}$ (ref. 19), *Cxcl12*^{DsRed} (ref. 18), *Cxcl12*$^{fl/+}$ (ref. 18), *R26*^{tdTomato} (ref. 26), *Vav1-cre* (ref. 24), *Lepr*^{cre} (ref. 27), *Tcf21*$^{cre/ER}$ (ref. 21) and α-*catulin*^{GFP}. To induce Cre/ER activity in *Tcf21*$^{cre/ER}$ mice, 4–6-week-old mice were administered 2 mg tamoxifen (Sigma) daily by oral gavage for 12 consecutive days. For induction of EMH, mice were injected at day 0 with a single dose of 4 mg cyclophosphamide followed by daily injections of 5 μg G-CSF for 4–21 days. Both male and female mice were used. All mice were housed in the Animal Resource Center at the University of Texas Southwestern Medical Center (UTSW). All procedures were approved by the UTSW Institutional Animal Care and Use Committee.

Flow cytometric analysis of haematopoietic cells. Bone marrow cells were isolated by flushing the femur or tibia with Ca^{2+}- and Mg^{2+}-free HBSS with 2% heat-inactivated bovine serum using a 3 ml syringe fitted with a 25-gauge needle. Spleen cells were obtained by crushing the spleen between two frosted slides. The cells were dissociated to a single-cell suspension by gently passing through the needle several times and then filtering through a 40-μm nylon mesh. Blood was collected by cardiac puncture, and white blood cells were isolated by ficoll centrifugation according to the manufacturer's instructions (GE Healthcare). The following antibodies were used to isolate HSCs: anti-CD150 (TC15-12F12.2), anti-CD48 (HM48-1), anti-Sca-1 (E13-161.7), anti-c-kit (2B8) and the following antibodies against lineage markers (anti-Ter119, anti-B220 (6B2), anti-Gr-1 (8C5), anti-CD2 (RM2-5), anti-CD3 (17A2), anti-CD5 (53-7.3) and anti-CD8 (53-6.7)). Haematopoietic progenitors were identified by flow cytometry using the following antibodies: anti-Sca-1 (E13-161.7), anti-c-Kit (2B8) and the following antibodies against lineage markers (anti-Ter119, anti-B220 (6B2), anti-Gr-1 (8C5), anti-CD2 (RM2-5), anti-CD3 (17A2), anti-CD5 (53-7.3) and anti-CD8 (53-6.7)), anti-CD34 (RAM34), anti-CD135 (Flt3) (A2F10), anti-CD16/32 (FcγR) (93), anti-CD127 (IL7Rα) (A7R34), anti-CD24 (M1/69), anti-CD43 (1B11), anti-B220 (6B2), anti-IgM (II/41), anti-CD3 (17A2), anti-Gr-1 (8C5), anti-Mac-1 (M1/70), anti-CD41 (MWReg30), anti-CD71 (C2) and anti-Ter119. 4′,6-Diamidino-2-phenylindole (DAPI) was used to exclude dead cells. Antibodies were obtained from eBioscience or BD Bioscience.

Flow cytometric analysis of stromal cells. To isolate bone marrow stromal cells the marrow was gently flushed out of the bone marrow cavity with a 3-ml syringe fitted with a 23-guage needle and then transferred into 1 ml pre-warmed bone marrow digestion solution (200 U ml^{-1} DNase I (Sigma), 250 μg ml^{-1} Liberase^{DL} (Roche) in HBSS plus Ca^{2+} and Mg^{2+}) and incubated at 37 °C for 30 min with gentle shaking. To isolate splenic stromal cells, the spleen capsule was cut into ~1 mm^3 fragments using scissors and then digested as described earlier in spleen digestion solution (200 U ml^{-1} DNase I, 250 μg ml^{-1} Liberase^{DL}, 1 mg ml^{-1} collagenase, type 4 (Roche) and 500 μg ml^{-1} collagenase D (Roche) in HBSS plus Ca^{2+} and Mg^{2+}). After a brief vortex, the spleen fragments were allowed to sediment for ~3 min and the supernatant was transferred to another tube on ice. The sedimented (undigested) spleen fragments were subjected to a second round of digestion. The two fractions of digested cells were pooled and filtered through a 100-μm nylon mesh. Anti-PDGFR-α (APA5), anti-PDGFR-β (APB5), anti-LepR (R&D), anti-CD45 (30F-11) and anti-Ter119 antibodies were used to isolate stromal cells. For analysis of endothelial cells, mice were injected intravenously into the retro-orbital venous sinus with 10 μg Alexa-Fluor-660-conjugated anti-VE-cadherin antibody (BV13) 10 min before being killed. Samples were analysed using a FACSAria or FACSCanto II flow cytometer (BD Biosciences).

BrdU incorporation assay. To assess BrdU incorporation into spleen cells after EMH induction, mice were intraperitoneally injected with a single dose of BrdU (2 mg BrdU per mouse) then maintained on 0.5 mg BrdU per ml drinking water for 7 days. Endothelial cells were labelled by intravenous injection of an anti-VE-cadherin antibody (eBioscience). Enzymatically dissociated spleen cells were stained with antibodies against surface markers and the target cell populations were sorted then resorted to ensure purity. The sorted cells were then fixed, and stained with an anti-BrdU antibody using the BrdU APC Flow Kit (BD Biosciences) according to the manufacturer's instructions.

Long-term competitive reconstitution assay. Adult recipient mice were irradiated using an XRAD 320 X-ray irradiator (Precision X-Ray) with two doses of 540 rad (total 1,080 rad) delivered at least 2 h apart. Cells were injected into the retro-orbital venous sinus of anaesthetized mice. Sorted doses of splenocytes from donor mice with EMH were transplanted along with 3×10^5 recipient bone marrow cells. Recipient mice were bled every 4 weeks to assess the level of donor-derived blood cells, including myeloid, B and T cells for at least 16 weeks. Blood was subjected to ammonium chloride/potassium red cell lysis before antibody staining. Antibodies including anti-CD45.2 (104), anti-CD45.1 (A20), anti-Gr1 (8C5), anti-Mac-1 (M1/70), anti-B220 (6B2) and anti-CD3 (KT31.1) were used for flow cytometric analysis.

Tissue sectioning and confocal imaging. For bone marrow sections, freshly dissected bones were fixed in 4% paraformaldehyde overnight followed by 3 days of decalcification in 10% EDTA dissolved in PBS. Bones were sectioned using the CryoJane tape-transfer system (Instrumedics). For spleen sections, freshly dissected spleens were fixed in 4% paraformaldehyde for 1 h followed by 1 day incubation in 10% sucrose in PBS. Frozen spleens were sectioned with a cryostat (Leica). For whole mount imaging, spleens were sectioned into ~2 mm pieces. Spleen sections were blocked in PBS with 10% horse serum for 1 h and then stained overnight with chicken-anti-GFP (Aves) and/or rabbit-anti-laminin (Abcam) antibodies. Donkey-anti-chicken Alexa Fluor 488 and/or donkey-anti-rabbit Alexa Fluor 647 were used as secondary antibodies (Invitrogen). Specimens were mounted with anti-fade prolong gold (Invitrogen) and images were acquired with either a Zeiss LSM780 confocal microscope or a Leica SP8 confocal microscope equipped with a resonant scanner. Three-dimensional images were achieved using Bitplane Imaris v.7.7.1 software.

Deep imaging of spleens. Spleens were harvested and fixed for 4 h in 4% PFA at 4 °C. Since the spleen capsule is highly autofluorescent, spleens were sectioned perpendicular to the long axis into 300-μm-thick sections using a Leica VT100S vibrotome. These 300-μm sections were fixed for an additional 2 h in 4% PFA and blocked overnight in staining solution (10% dimethylsulfoxide (DMSO), 0.5% IgePal630 (Sigma) and 5% donkey serum (Jackson Immunoresearch) in PBS). All staining steps were performed in staining solution on a rotator at room temperature. Spleen sections were stained for 3 days in primary antibodies, washed overnight in several changes of PBS then stained for 3 days in secondary antibodies. The stained sections were dehydrated in a methanol dehydration series then incubated for 3 h in 100% methanol with several changes. The methanol was then exchanged with benzyl alcohol:benzyl benzoate 1:2 mix (BABB clearing[28]). The tissues were incubated in BABB for 3 h to overnight with several exchanges of fresh BABB. Spleen sections were mounted in BABB between two coverslips and sealed with silicone (Premium waterproof silicone II clear; General Electric). We found it necessary to clean the BABB of peroxides (which can accumulate as a result of exposure to air and light) by adding 10 g of activated aluminium oxide (Sigma) to 40 ml of BABB and rotating for at least 1 h, then centrifuging at 2,000 *g* for 10 min to remove the suspended aluminium oxide particles. Images were acquired using a Zeiss LSM780 confocal microscope with a Zeiss LD LCI Plan-Apo ×25/0.8 multi-immersion objective lens, which has a 570 μm working distance. Images were taken at 512 × 512 pixel resolution with 2 μm Z-steps, pinhole for the internal detector at 47.7 μm. Random spots were inserted into images by generating randomized *X*, *Y*, and *Z* coordinates using the random integer generator at http:// www.random.org.

Splenectomy. After mouse anaesthesia by ketamine/xylazine, a ventral midline incision was made and the peritoneum was breached. The splenic blood vessels were ligated with an absorbable suture (4-0 vicryl). The splenic vessels were cut distal to the suture and the spleen was removed. The vessels were cauterized and the abdomen was sutured with non-absorbable sutures (3-0 Tevdek III). Buprenorphine was administered every 12 h for 3 days to minimize postoperative pain and mice were maintained with ampicillin-containing water to avoid infection. Complete blood counts were measured one month after the survival surgery.

Induction of EMH by bleeding. EMH was induced by repeated bleeding over a 2-week period according to a published protocol[2]. Briefly, 4–6 month-old mice were bled via the tail vein five times, every 3 days, removing approximately 250 μl of blood each time, then the mice were killed for analysis 2 days after the last bleed.

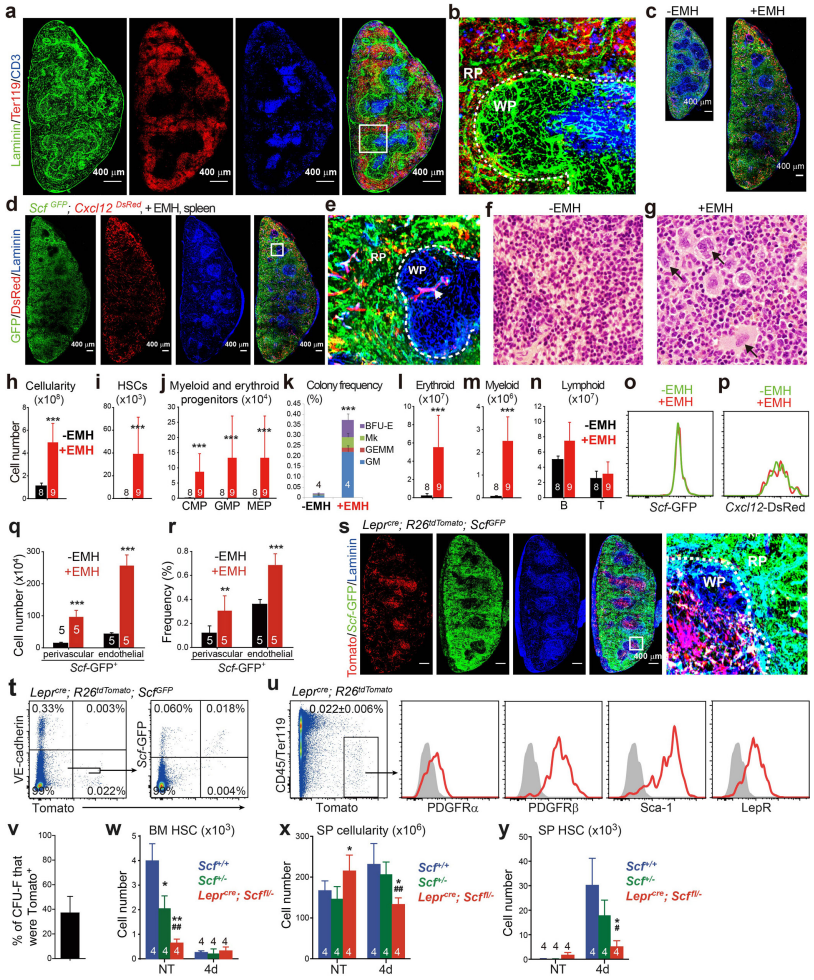
Western blot. Approximately 30,000 CD45$^-$Ter119$^-$VE-cadherin$^+$ splenic endothelial cells were flow cytometrically sorted into 50 μl of 66% trichoracetic acid (TCA) in water. Extracts were incubated on ice for at least 15 min and centrifuged at 16,100 *g* at 4 °C for 10 min. Precipitates were washed in acetone twice and the dried pellets were solubilized in 9 M urea, 2% Triton X-100, and 1% dithiothreitol (DTT). Samples were separated on 4–12% Bis-Tris polyacrylamide gels (Invitrogen) and transferred to PVDF membrane (Millipore). The blots were incubated with primary antibodies overnight at 4 °C and then with secondary antibodies. Blots were developed with the SuperSignal West Femtochemiluminescence kit (Thermo Scientific). Primary antibodies used: rabbit-anti-SCF (Abcam, 1:1,000) and mouse-anti-actin (Santa Cruz, clone AC-15, 1:20,000).

Quantitative real-time PCR. Cells were sorted directly into Trizol (Life Technologies). Total RNA was extracted according to the manufacturer's instructions. Total RNA was reverse transcribed using SuperScript III Reverse Transcriptase (Life Technologies). Quantitative real-time PCR was performed using SYBR green on a LightCycler 480 (Roche). β-Actin was used to normalize the RNA content of samples. Primers used in this study were *Scf*: 5′-GCCAGAAACTAGATCCTTTACTCCTGA-3′ and 5′-CATAAATGGTTTTGTGACACTGACTCTG-3′; β-actin: 5′-GCTCTTTTCCAGCCTTCCTT-3′ and 5′-CTTCTGCATCCTGTCAGCAA-3′.

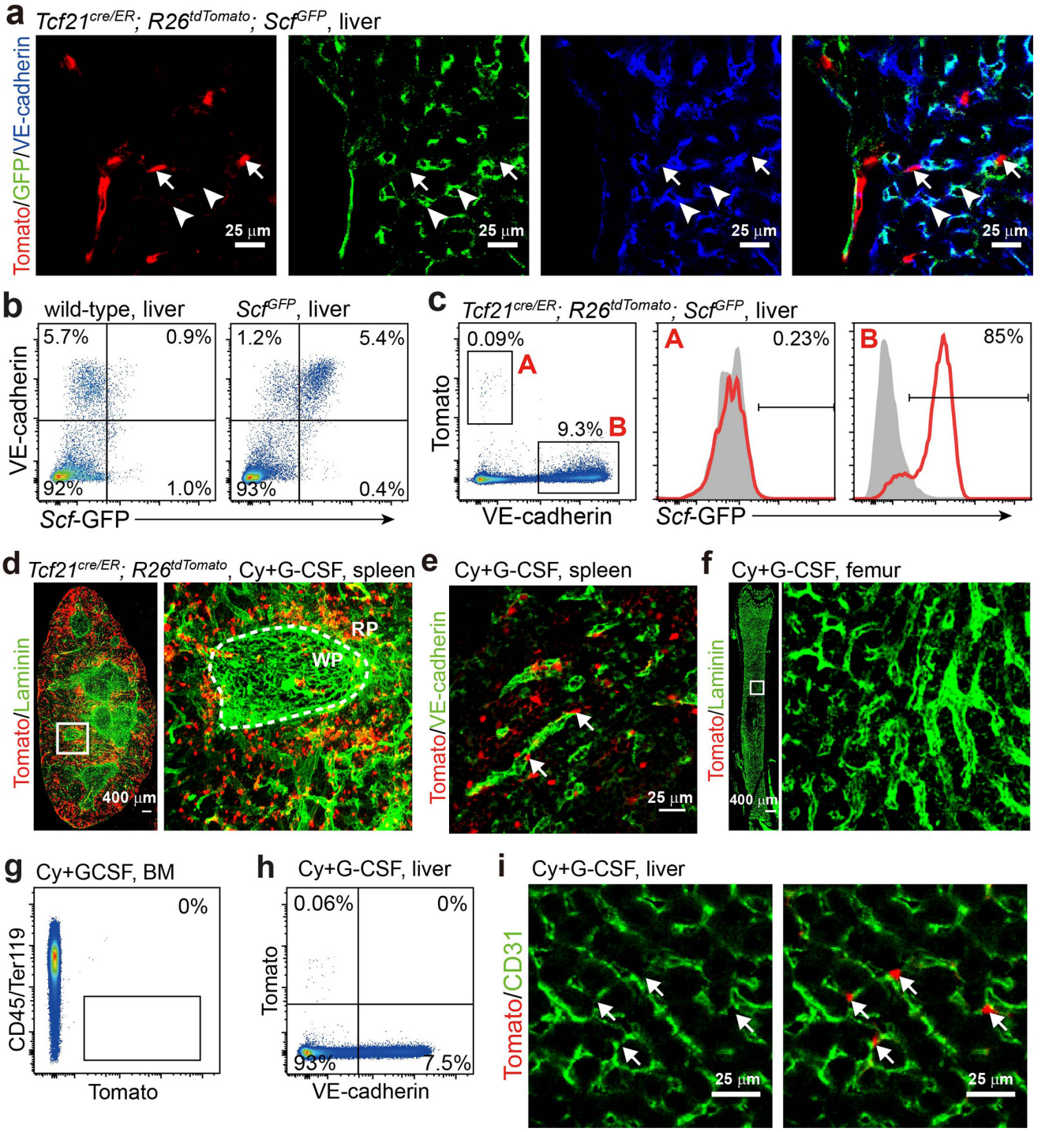
Gene expression profiling. Three independent samples of 5,000 spleen *Scf*-GFP^{+}VE-cadherin^{-} spleen stromal cells and two independent samples of 5,000 unfractionated spleen cells were flow cytometrically sorted into Trizol. Total RNA was extracted, amplified, and sense strand cDNA was generated using the Ovation Pico WTA System V2 (NuGEN) according to the manufacturer's instructions. cDNA was fragmented and biotinylated using the Encore Biotin Module (NuGEN) according to the manufacturer's instructions. Labelled cDNA was hybridized to Affymetrix Mouse Gene ST 1.0 chips according to the manufacturer's instructions. Expression values for all probes were normalized and determined using the robust multi-array average (RMA) method[29].

Statistical methods. Panels in all figures represented multiple independent experiments performed on different days with different mice. Sample sizes were not based on power calculations. No randomization or blinding was performed. No animals were excluded from analysis. Variation is always indicated using standard deviation. For analysis of the statistical significance of differences between two groups we generally performed two-tailed Student's *t*-tests. For analysis of the statistical significance of differences among more than two groups, we performed repeated measures one-way analysis of variance (ANOVA) tests with Greenhouse–Geisser correction (variances between groups were not equal) and Tukey's multiple comparison tests with individual variances computed for each comparison. To assess the statistical significance of differences in fetal mass between paired control and mutant mice (Fig. 5j and Extended Data Fig. 8v), we performed a two-way ANOVA.

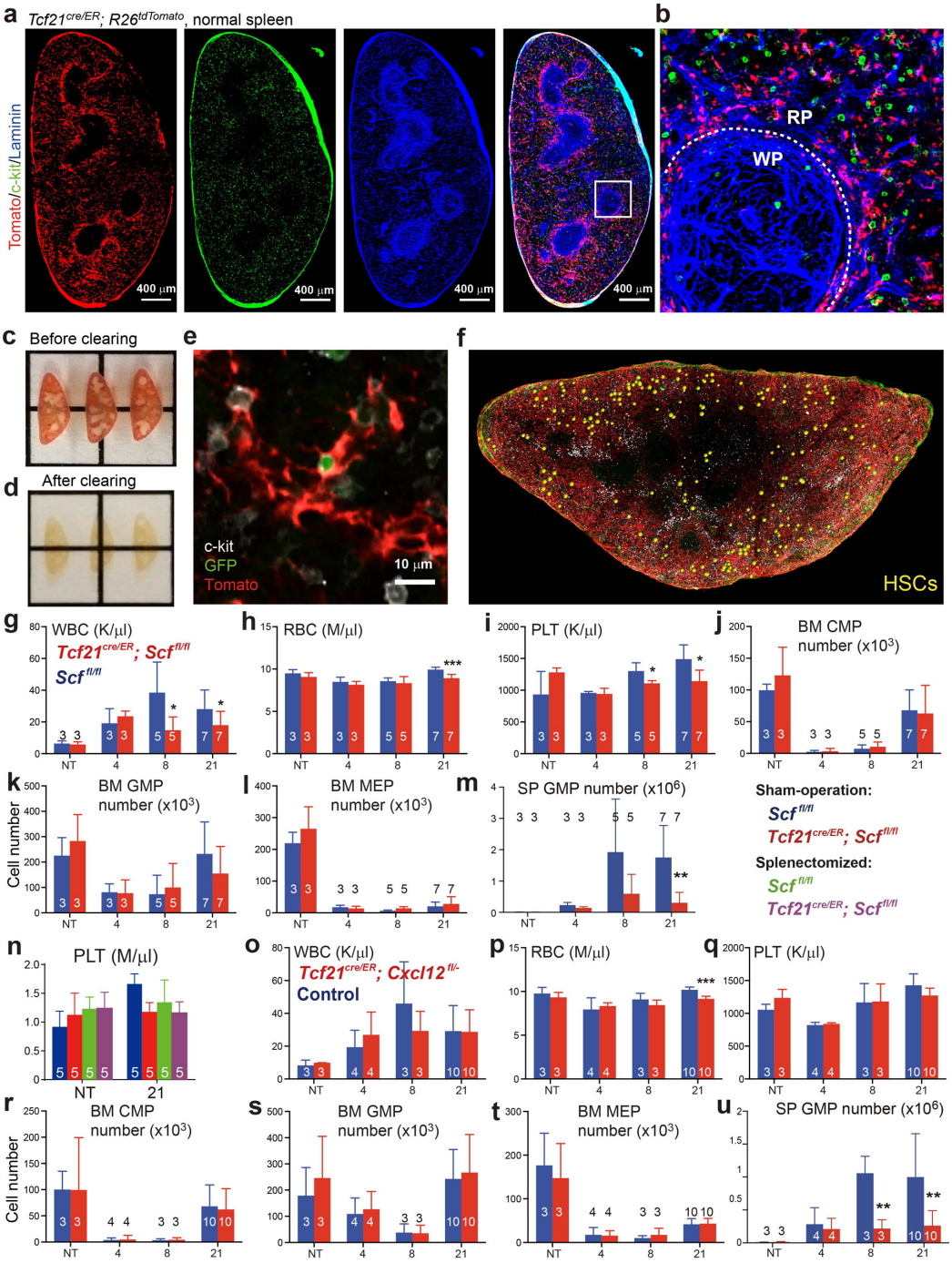
26. Madisen, L. *et al.* A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nature Neurosci.* **13,** 133–140 (2010).
27. DeFalco, J. *et al.* Virus-assisted mapping of neural inputs to a feeding center in the hypothalamus. *Science* **291,** 2608–2613 (2001).
28. Becker, K., Jährling, N., Saghafi, S. & Dodt, H. U. Immunostaining, dehydration, and clearing of mouse embryos for ultramicroscopy. *Cold Spring Harb. Protoc.* **2013,** 743–744 (2013).
29. Irizarry, R. A. *et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4,** 249–264 (2003).
30. Cesta, M. F. Normal structure, function, and histology of the spleen. *Toxicol. Pathol.* **34,** 455–465 (2006).
31. Tronche, F. *et al.* Disruption of the glucocorticoid receptor gene in the nervous system results in reduced anxiety. *Nature Genet.* **23,** 99–103 (1999).
32. Zhu, X., Bergles, D. E. & Nishiyama, A. NG2 cells generate both oligodendrocytes and gray matter astrocytes. *Development* **135,** 145–157 (2008).
33. Zhu, X. *et al.* Age-dependent fate and lineage restriction of single NG2 cells. *Development* **138,** 745–753 (2011).
34. Logan, M. *et al.* Expression of Cre recombinase in the developing mouse limb bud driven by a Prxl enhancer. *Genesis* **33,** 77–80 (2002).
35. Rivers, L. E. *et al.* PDGFRA/NG2 glia generate myelinating oligodendrocytes and piriform projection neurons in adult mice. *Nature Neurosci.* **11,** 1392–1401 (2008).
36. Cuttler, A. S. *et al.* Characterization of Pdgfrb-Cre transgenic mice reveals reduction of ROSA26 reporter activity in remodeling arteries. *Genesis* **49,** 673–680 (2011).
37. Holtwick, R. *et al.* Smooth muscle-selective deletion of guanylyl cyclase-A prevents the acute but not chronic effects of ANP on blood pressure. *Proc. Natl Acad. Sci. USA* **99,** 7142–7147 (2002).
38. Xin, H. B., Deng, K. Y., Rishniw, M., Ji, G. & Kotlikoff, M. I. Smooth muscle expression of Cre recombinase and eGFP in transgenic mice. *Physiol. Genomics* **10,** 211–215 (2002).
39. Wendling, O., Bornert, J. M., Chambon, P. & Metzger, D. Efficient temporally-controlled targeted mutagenesis in smooth muscle cells of the adult mouse. *Genesis* **47,** 14–18 (2009).



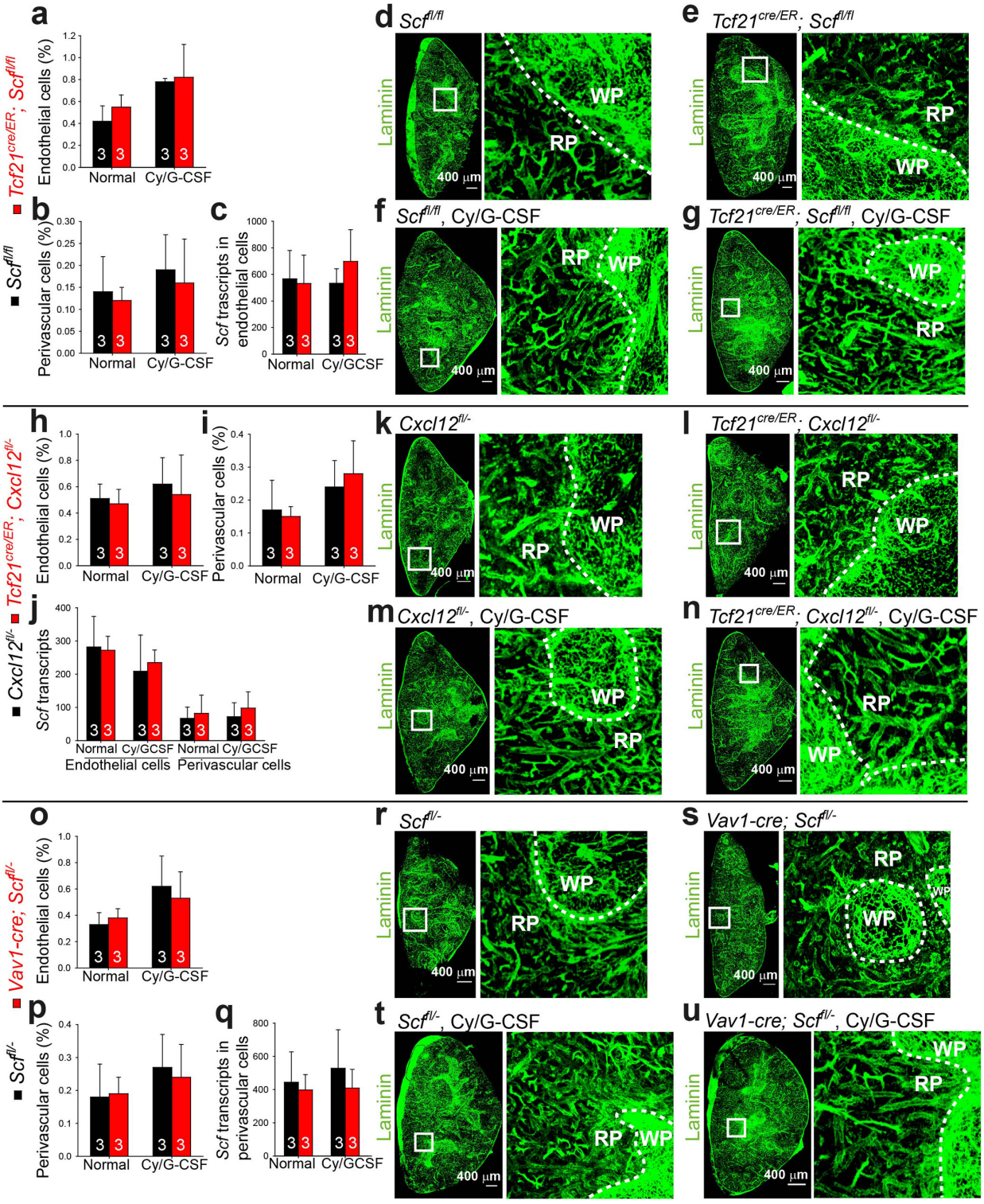
Extended Data Figure 1 | Cy+21 d G-CSF treatment induces EMH in the spleen and deletion of *Scf* from LepR$^+$ cells significantly reduces the number of HSCs in the bone marrow and the spleen after induction of EMH. a,**b**, Staining with anti-laminin antibody distinguished the vasculature of red pulp (RP) from white pulp (WP). The red pulp and white pulp were marked by clusters of Ter119$^+$ cells (red) and CD3$^+$ cells (blue), respectively[30]. Dashed line depicts the boundary between red pulp and white pulp (representative images from 3 mice in 3 independent experiments). **c**, Spleen sections of the same magnification show the enlargement of the spleen after induction of EMH by Cy+21 d G-CSF. These are the same images as in Fig. 1a,d, adjusted to reflect the same magnification. **d**, **e**, Imaging of thick spleen sections from *Scf^{GFP}; Cxcl12^{DsRed}* mice after the induction of EMH by Cy+21 d G-CSF. **e**, High-magnification view of the boxed area in **d**. Dashed lines depict the boundaries between white pulp and red pulp. Arrow indicates the central arteriole in the white pulp around which stromal cells expressed *Cxcl12*-DsRed (representative images from 3 mice from 3 independent experiments). **f**, **g**, Haematoxylin and eosin (H&E) staining showing the increase in haematopoiesis in the spleen after induction of EMH using Cy+G-CSF (+EMH, **g**) as evidenced by the presence of megakaryocytes (arrows; $n = 3$ mice per condition from 3 independent experiments). **h**–**n**, Cy+G-CSF treatment significantly increased spleen cellularity (**h**), as well as the numbers of HSCs (**i**), MEPs (**j**), frequencies of colony-forming progenitors (**k**), numbers of Ter119$^+$ erythroid cells (**l**) and Gr-1$^+$Mac-1$^+$ myeloid cells (**m**) in the spleen but not the number of B220$^+$ or CD3$^+$ lymphoid cells (**n**). The numbers of mice per treatment are shown in each bar of each panel. Each panel shows mean ± s.d. from five independent experiments. **o**, **p**, *Scf*-GFP (**o**) and *Cxcl12*-DsRed (**p**) fluorescence by spleen stromal cells before (−EMH) and after induction of EMH (+EMH) using Cy+G-CSF. **q**, **r**, The frequencies (**q**) and absolute numbers (**r**) of *Scf*-GFP$^+$VE-cadherin$^+$ endothelial cells and *Scf*-GFP$^+$VE-cadherin$^-$ stromal cells significantly increased upon induction of EMH by Cy+21 d G-CSF (+EMH). **s**, Spleens from *Lepr^{cre}; R26^{tdTomato}; Scf^{GFP}* mice showed Tomato expression was primarily in the stromal cells of the white pulp. Although most *Scf*-GFP expression was in endothelial cells and perivascular stromal cells of the red pulp (Fig. 1a–d), some *Scf*-GFP$^+$ stromal cells were in the white pulp, most of which appeared to express LepR. Dashed line depicts the boundary between red pulp and white pulp (representative images of 6 mice from 4 independent experiments). **t**, Flow cytometric analysis of enzymatically dissociated spleen cells from *Lepr^{cre}; R26^{tdTomato}; Scf^{GFP}* mice showed that only a small minority of non-endothelial *Scf*-GFP$^+$ cells were positive for Tomato ($n = 3$ mice from 3 independent experiments). **u**, Tomato$^+$CD45$^-$Ter119$^-$ stromal cells in the spleens of *Lepr-cre; R26^{tdTomato}* mice expressed PDGFR-α, PDGFR-β, Sca-1 and LepR ($n = 3$ mice from 3 independent experiments). **v**, Percentage of all CFU-F colonies formed by enzymatically dissociated spleen cells from *Lepr^{cre}; R26^{tdTomato}* mice that expressed Tomato. Macrophage colonies were excluded by staining with anti-CD45 antibody ($n = 4$ mice from 3 independent experiments). **w**, *Lepr^{cre}; Scf$^{fl/-}$* mice had significantly fewer HSCs in the bone marrow than wild-type and *Scf$^{fl/-}$* controls before induction of EMH ($n = 4$ mice per genotype per time point mice from 4 independent experiments). NT, not treated. **x**, **y**, *Lepr^{cre}; Scf$^{fl/-}$* mice displayed significantly lower spleen cellularity (**x**) and HSC number (**y**) in the spleen than wild-type and *Scf$^{fl/-}$* controls after induction of EMH with cyclophosphamide plus 4 days of G-CSF. The numbers of mice per treatment are shown in each bar. Data represent mean ± s.d. from 4 independent experiments. **h**–**n**, **q**, **r**, The statistical significance of differences was assessed using two-tailed Student's *t*-tests ($***P < 0.001$). **w**–**y**, The statistical significance of differences between genotypes was assessed using repeated measures one-way ANOVAs with Greenhouse–Geisser correction and Tukey's multiple comparison tests with individual variances computed for each comparison. $*P < 0.05$, $**P < 0.01$, statistical significance relative to wild-type (*Scf$^{+/+}$*). $\dagger P < 0.05$, $\dagger\dagger P < 0.01$, statistical significance between *Scf$^{+/-}$* and *Lepr^{cre}; Scf$^{fl/-}$*.



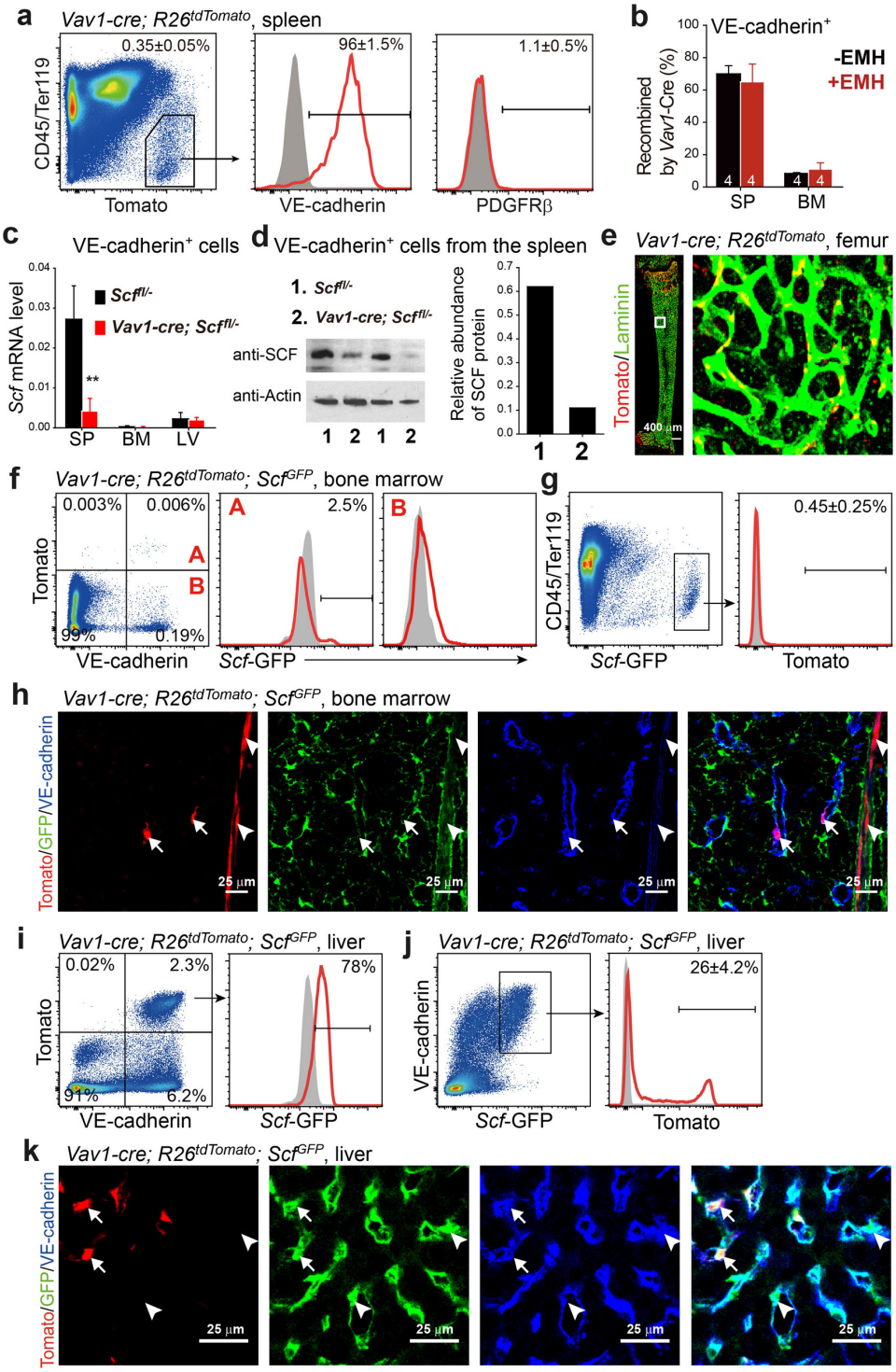
Extended Data Figure 2 | *Scf* is expressed by most endothelial cells but not by *Tcf21*$^+$ perivascular cells in the liver; Cy+21 d G-CSF does not significantly change the recombination pattern of *Tcf21*-Cre/ER in the spleen, bone marrow or liver. a–i, To identify Cre alleles that recombine in spleen, but not bone marrow, stromal cells we assessed the gene expression profile of spleen *Scf*-GFP$^+$VE-cadherin$^-$ stromal cells (Extended Data Table 1). Nestin, *NG2* (also known as *Cspg4*) and *Prx1* were low or undetectable (data not shown). *Nestin*-Cre[31], *NG2*-Cre[32], *NG2*-Cre/ER[33], and *Prx1*-Cre[34] did not recombine widely or specifically in *Scf*-GFP$^+$ stromal cells in the spleen (data not shown). *Pdgfra* and *Pdgfrb* were expressed by spleen *Scf*-GFP$^+$ stromal cells but neither *Pdgfra*-Cre/ER (ref. 35) nor *Pdgfrb*-Cre (ref. 36) recombined efficiently (data not shown). *Sm22* (also known as *Tagln*), *Myh11*, *Sma* (also known as *Acta2*) and *Tcf21* were significantly more highly expressed by spleen than bone marrow *Scf*-GFP$^+$ stromal cells (Extended Data Table 1). *Sm22*-Cre (ref. 37), *Myh11*-Cre (ref. 38) and *Sma*-Cre/ER (ref. 39) recombined in few spleen *Scf*-GFP$^+$ stromal cells (data not shown). However, *Tcf21*-Cre/ER recombined in perivascular stromal cells in the spleen but not bone marrow (Fig. 2). **a–c**, Under normal conditions, *Scf*-GFP was expressed by most VE-cadherin$^+$ endothelial cells (arrowheads in **a**) but not by *Tcf21*$^+$ stromal cells (arrows in **a**) in the liver ($n=3$ mice from 3 independent experiments). **d**, **e**, EMH induced by Cy+21 d G-CSF did not alter the general distribution (**d**) or perivascular localization (**e**) of Tomato$^+$ cells in the spleens of *Tcf21*$^{cre/ER}$; *R26*^{tdTomato} mice as compared to normal mice (Fig. 2a, d). **f**, **g**, Tomato expression was undetectable in the bone marrow of *Tcf21*$^{cre/ER}$; *R26*^{tdTomato} mice after Cy+G-CSF treatment irrespective of whether the bone marrow was analysed by whole-mount imaging (**f**) or flow cytometry (**g**). **h**, **i**, EMH induced by Cy+G-CSF did not significantly change the frequency (**h**) or perivascular localization (**i**, arrows) of Tomato$^+$ cells in the livers of *Tcf21*$^{cre/ER}$; *R26*^{tdTomato} mice. **d–i**, $n=3$ mice from 3 independent experiments.



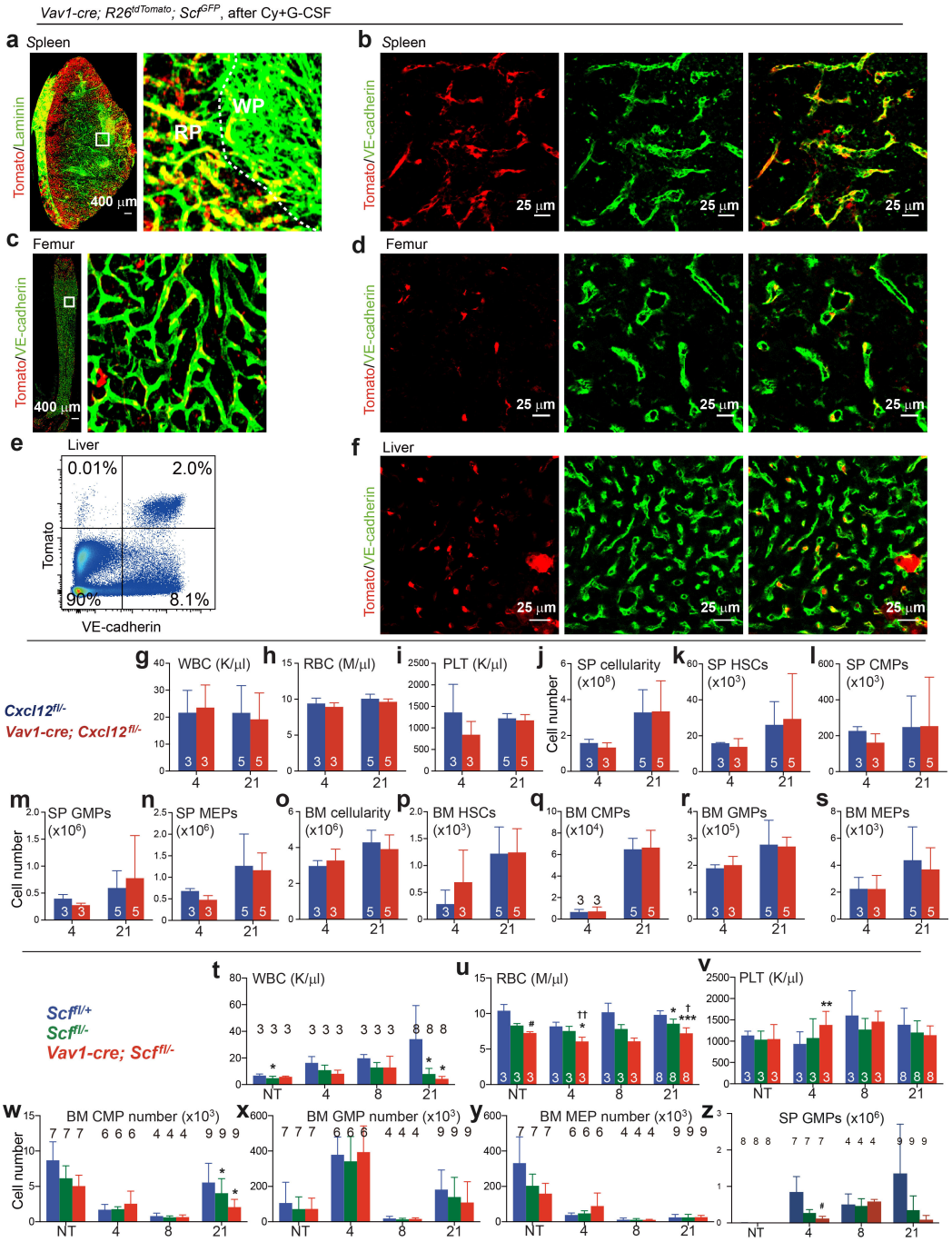
Extended Data Figure 3 | Deep imaging of HSCs in the spleen; deletion of *Scf* or *Cxcl12* from *Tcf21*-expressing stromal cells in the spleen reduced peripheral blood cell counts but did not affect bone marrow haematopoiesis. a, **b**, The vast majority of c-Kit$^+$ haematopoietic progenitors localized adjacent to *Tcf21*-expressing stromal cells in the red pulp of the normal spleen ($n = 3$ mice from 3 independent experiments). **c**, **d**, Three-hundred-micrometre-thick sections of spleen before (**c**) and after optical clearing (**d**). **e**, **f**, Deep imaging of α-*catulin*-GFP$^+$c-Kit$^+$ HSCs in cleared spleen segments from *Tcf21*$^{cre/ER}$; *R26*^{tdTomato}; α-*catulin*^{GFP} mice. A representative high-magnification image of an α-*catulin*-GFP$^+$c-Kit$^+$ HSC surrounded by Tomato$^+$ stromal cells (**e**). **f**, Low-magnification view of a digitally reconstructed 300-μm-thick spleen fragment with α-*catulin*-GFP$^+$c-Kit$^+$ HSCs identified by large yellow spheres. Note that actual HSCs would be smaller than the yellow spheres but would not be visible at this magnification ($n = 3$ mice from 3 independent experiments). **g**–**m**, *Tcf21*$^{cre/ER}$; *Scf*$^{fl/fl}$ and *Scf*$^{fl/fl}$ control mice were treated with tamoxifen then examined 1 month later without further treatment (not treated (NT)) or after treatment with cyclophosphamide plus 4, 8, or 21 days of G-CSF to induce EMH. Data show WBC (**g**), RBC (**h**) and PLT counts (**i**), numbers of CMPs (**j**), GMPs (**k**) and MEPs (**l**) in the bone marrow and numbers of GMPs in the spleen (**m**). **n**, Platelet counts of sham-operated and splenectomized mice that were treated with Cy+21 d G-CSF 1 month after surgery. **o**–**u**, *Tcf21*$^{cre/ER}$; *Cxcl12*$^{fl/-}$ mice and littermate controls (*Cxcl12*$^{fl/-}$ or *Cxcl12*$^{+/-}$) were treated with tamoxifen then examined 1 month later without further treatment (NT) or after treatment with cyclophosphamide plus 4, 8, or 21 days of G-CSF to induce EMH. Data show WBC (**o**), RBC (**p**) and PLT counts (**q**), numbers of CMPs (**r**), GMPs (**s**) and MEPs (**t**) in the bone marrow and numbers of GMPs in the spleen (**u**). The numbers of mice per treatment are shown in each panel. All data reflect mean ± s.d. from 3 independent experiments. Two-tailed Student's *t*-tests were used to assess statistical significance (*$P < 0.05$, ***$P < 0.001$).



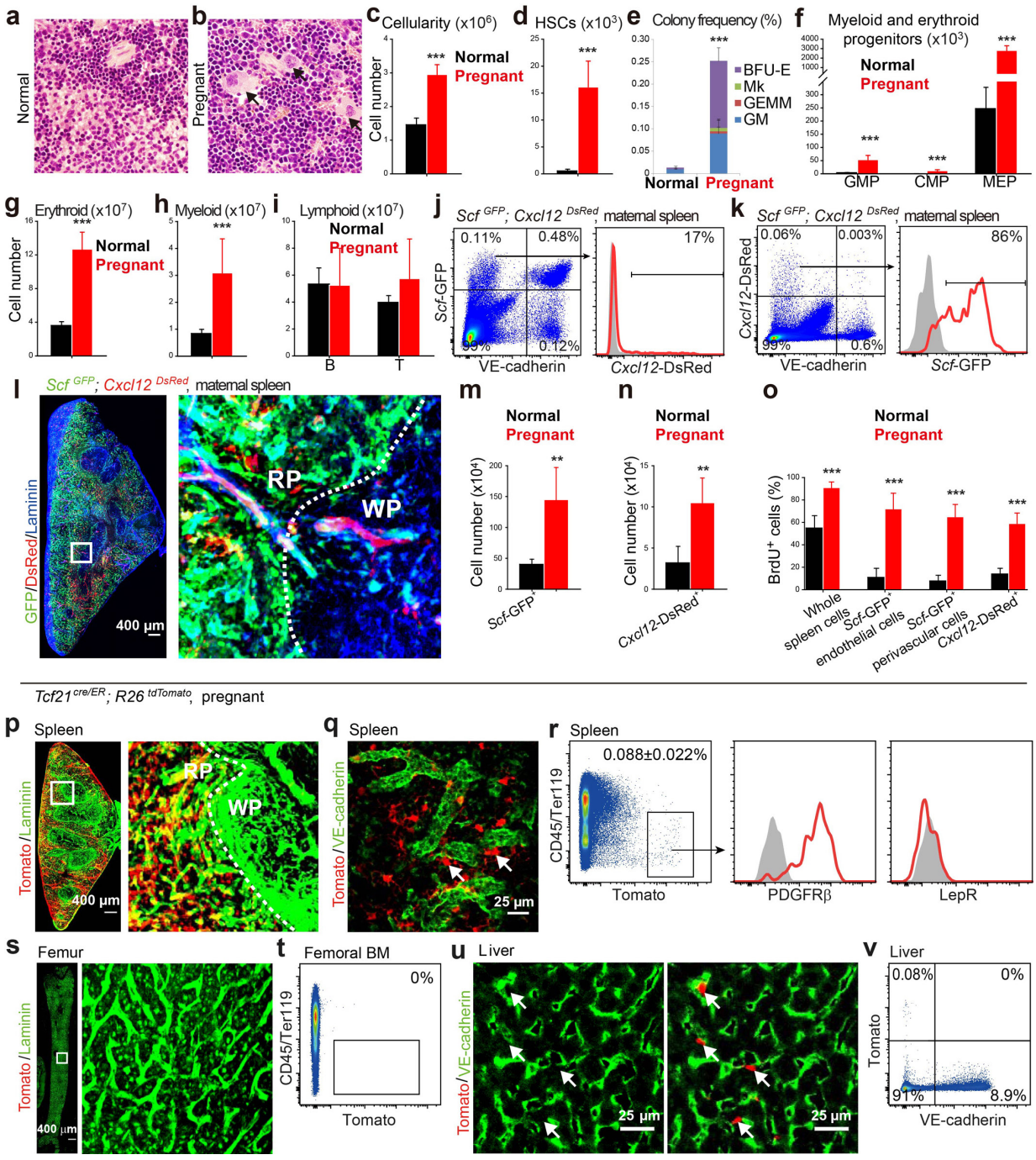
Extended Data Figure 4 | Conditional deletion of *Scf* or *Cxcl12* with *Tcf21*-Cre/ER, or *Scf* with *Vav1*-Cre, does not significantly affect the frequency or morphology of stromal cells in the spleen, irrespective of EMH induction. a–g, Irrespective of whether the mice were treated with Cy+21 d G-CSF, conditional deletion of *Scf* from *Tcf21*$^+$ cells did not significantly change the frequency of VE-cadherin$^+$ endothelial cells (**a**) or PDGFR-β$^+$ perivascular stromal cells (**b**), *Scf* transcript levels in endothelial cells (**c**), or the morphology or density of blood vessels in the spleen (**d–g**). **h–n**, Irrespective of whether the mice were treated with Cy+G-CSF, conditional deletion of *Cxcl12* from *Tcf21*$^+$ cells did not significantly change the frequency of VE-cadherin$^+$ endothelial cells (**h**) or PDGFR-β$^+$ perivascular stromal cells (**i**), *Scf* transcript levels in endothelial cells or perivascular stromal cells (**j**), or the morphology or density of blood vessels in the spleen (**k–n**). **o–u**, Irrespective of whether the mice were treated with Cy+G-CSF, conditional deletion of *Scf* using *Vav1*-Cre did not significantly change the frequency of VE-cadherin$^+$ endothelial cells (**o**) or PDGFR-β$^+$ perivascular stromal cells (**p**), *Scf* transcript levels in perivascular stromal cells (**q**), or the morphology or density of blood vessels in the spleen (**r–u**). *Scf* transcript levels in flow cytometrically isolated cells were normalized to β-actin and then compared to whole spleen cells (**c**, **j** and **q**). The data reflect mean ± s.d. from 3 mice per genotype per condition in 3 independent experiments. Two-tailed Student's *t*-tests were used to assess statistical significance.



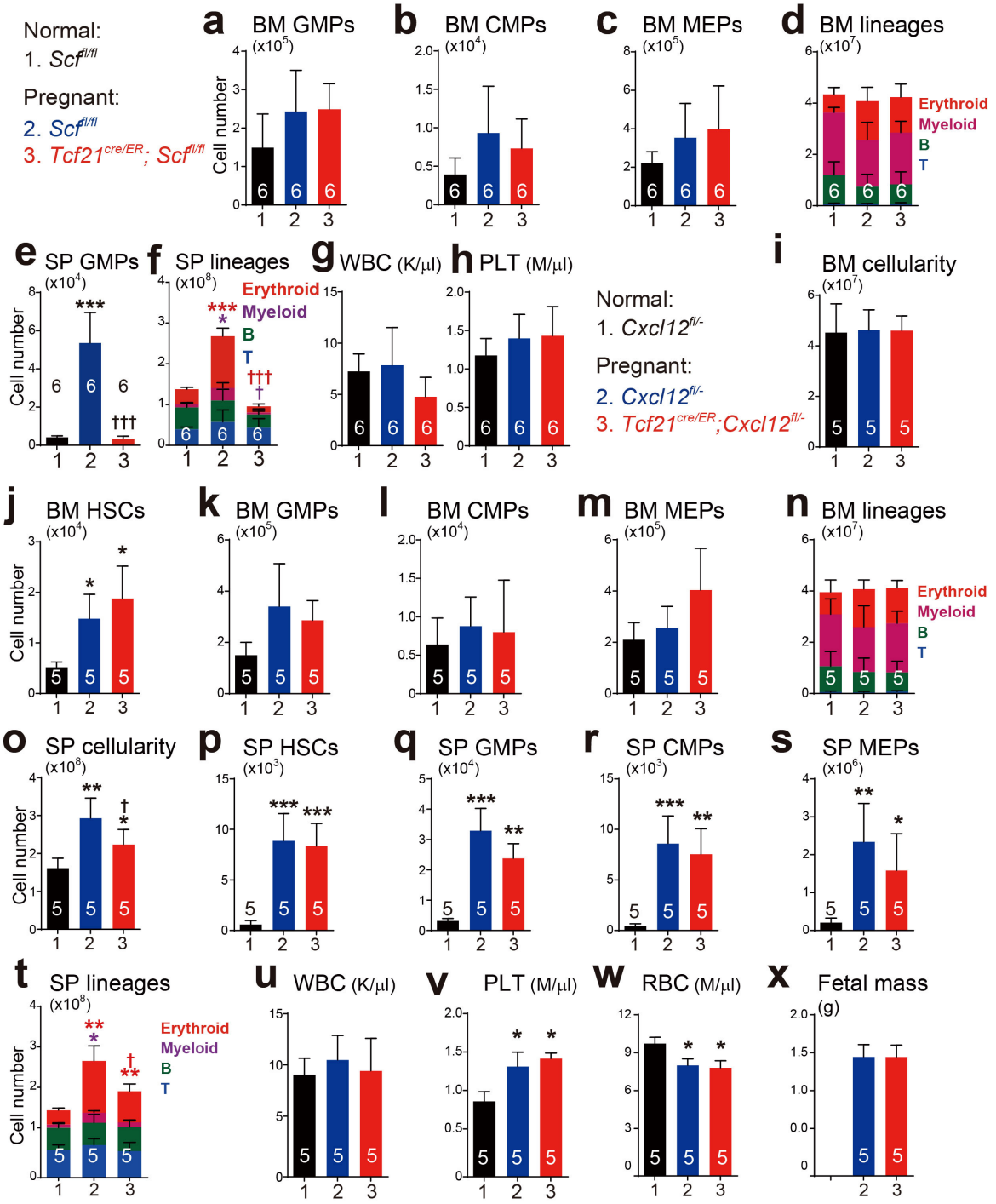
Extended Data Figure 5 | *Vav1*-Cre recombines efficiently and specifically in spleen endothelial cells but poorly in bone marrow or liver endothelial cells. a, Tomato^{high}CD45^{-}Ter119^{-} cells in *Vav1-cre; tdTomato* mice were uniformly positive for VE-cadherin and negative for PDGFR-β ($n = 3$ mice from 3 independent experiments). **b**, *Vav1*-Cre recombined in most spleen endothelial cells but in few bone marrow endothelial cells, irrespective of Cy+G-CSF treatment (+EMH). **c**, *Scf* transcript levels were significantly reduced in endothelial cells from the spleen but not from bone marrow or liver in *Vav1-cre; Scf$^{fl/-}$* mice as compared to *Scf$^{fl/-}$* mice. The *Scf* transcript level was normalized to β-actin. **d**, Western blot showed lower SCF protein levels in splenic endothelial cells from *Vav1-cre; Scf$^{fl/-}$* mice as compared to *Scf$^{fl/-}$* mice. SCF abundance was assessed relative to actin by Image J software ($n = 3$ mice per genotype from 3 independent experiments). **e**–**h**, In the bone marrow *Vav1*-Cre recombined in a minority of endothelial cells, including some sinusoidal (arrows in **h**) and some arteriolar (arrowheads in **h**) endothelial cells, that expressed little *Scf*-GFP by flow cytometry (**f**, **g**). The data reflect mean ± s.d. from 3 mice per genotype in 3 independent experiments. **i**–**k**, *Vav1*-Cre recombined inefficiently in liver endothelial cells. Most Tomato^{+} cells in the liver of *Vav1-cre; R26^{tdTomato}; Scf^{GFP}* mice were VE-cadherin^{+} and *Scf*-GFP^{+} (**i**; arrows in **k**) but these cells accounted for only 26 ± 4.2% of *Scf*-GFP^{+} cells by flow cytometry (**i**, **j**) and confocal microscopy (**k**, $n = 3$ mice from 3 independent experiments). Two-tailed Student's *t*-tests were used to assess statistical significance.



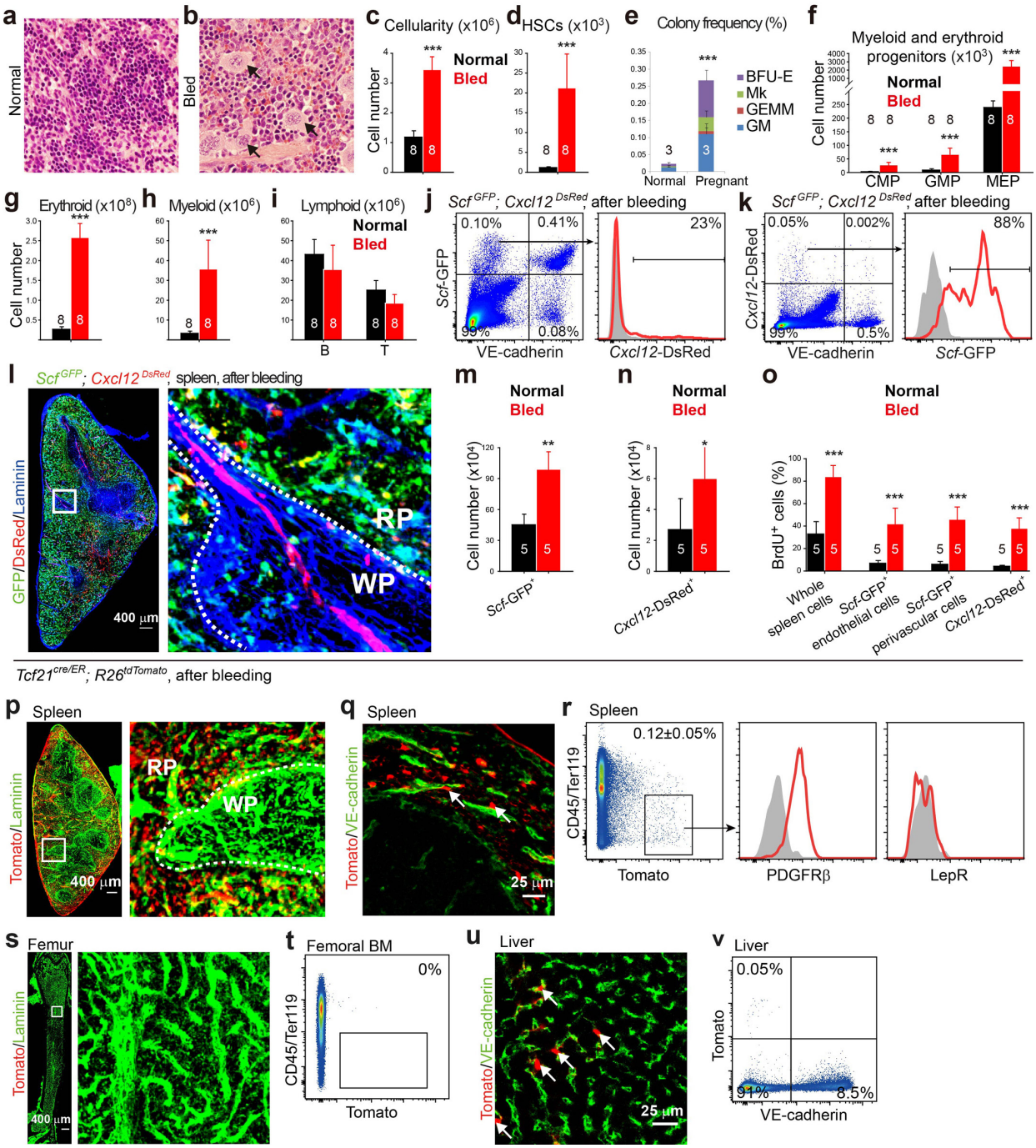
Extended Data Figure 6 | EMH induced by Cy+G-CSF does not significantly change the recombination pattern of *Vav1*-Cre in the spleen, bone marrow, or liver but deletion of *Scf* from endothelial cells in spleens with EMH reduces blood cell counts without affecting bone marrow haematopoiesis. a, **b**, After EMH induced by Cy+21 d G-CSF, *Vav1*-Cre-recombined cells were predominantly in the red pulp (**a**) and co-localized with VE-cadherin$^+$ cells (**b**) in the spleen. **c–f**, After EMH induced by Cy+21 d G-CSF, *Vav1*-Cre-recombined cells remained rare in the bone marrow (**c**, **d**) and liver (**e**, **f**; $n = 3$ mice from 3 independent experiments). **g–s**, *Vav1-cre; Cxcl12*$^{fl/-}$ mice and *Cxcl12*$^{fl/-}$ controls were treated with Cy+4–21 d G-CSF to induce EMH. Data show WBC (**g**), RBC (**h**), and platelet (**i**) counts, spleen cellularity (**j**) and numbers of HSCs (**k**), CMPs (**l**), GMPs (**m**) and MEPs (**n**) in the spleen as well as bone marrow cellularity (**o**), and numbers of HSCs (**p**), CMPs (**q**), GMPs (**r**) and MEPs (**s**) in one femur and one tibia. The data represent mean mean ± s.d. from 3 (Cy+4 d G-CSF treatment) and 5 (Cy+21 d G-CSF treatment) independent experiments. The number of mice per treatment is indicated on each bar. Two-tailed Student's *t*-tests were used to assess statistical significance. **t–z**, *Vav1-cre*; *Scf*$^{fl/-}$ mice and *Scf*$^{fl/+}$, *Scf*$^{fl/-}$ controls were treated with Cy+4–21 d G-CSF to induce EMH. Data show WBC (**t**), RBC (**u**), and platelet (PLT) (**v**) counts, numbers of CMPs (**w**), GMPs (**x**) and MEPs (**y**) in the bone marrow as well as numbers of GMPs in the spleen (**z**). Note that after 21 days of G-CSF both *Scf*$^{fl/-}$ and *Vav1-cre*; *Scf*$^{fl/-}$ mice showed significantly lower CMP numbers relative to *Scf*$^{fl/+}$ mice but their CMP numbers were not significantly different from each other (**w**), indicating that CMP numbers in the bone marrow were not influenced by *Scf* deletion from spleen endothelial cells. The data represent mean ± s.d. from 3 (no treatment (NT)), 3 (4 days), 3 (8 days), and 8 (21 days) independent experiments. The number of mice per treatment is indicated on each bar. The statistical significance of differences among genotypes was assessed using repeated measures one-way ANOVAs with Greenhouse–Geisser correction and Tukey's multiple comparison tests with individual variances computed for each comparison. $*P < 0.05$, $**P < 0.01$, statistical significance relative to *Scf*$^{fl/+}$ controls. $\dagger P < 0.05$, $\dagger\dagger P < 0.01$, statistical significance between *Scf*$^{fl/-}$ and *Vav1-cre*; *Scf*$^{fl/-}$.



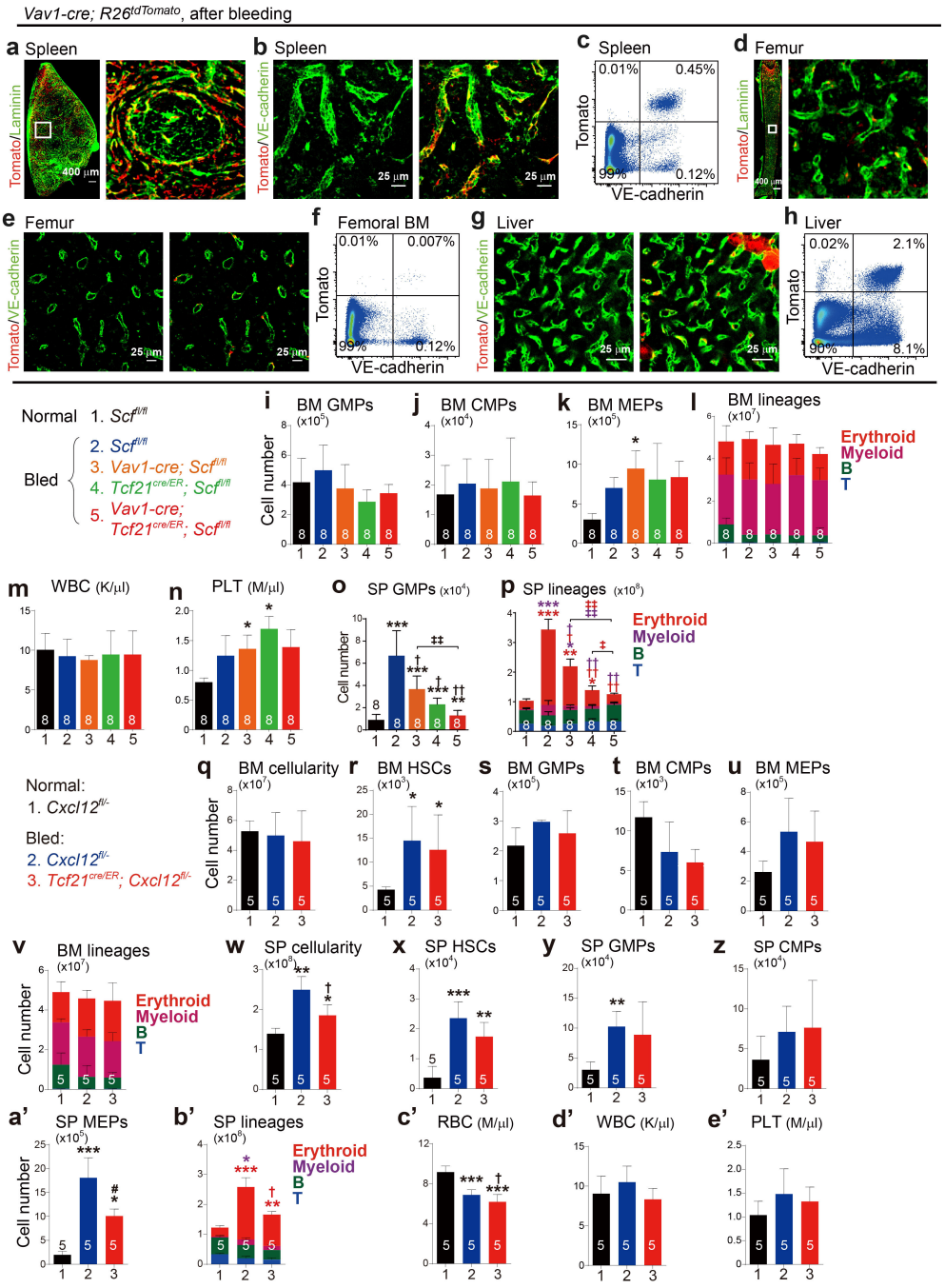
Extended Data Figure 7 | Pregnancy induces EMH and the proliferation of endothelial cells and stromal cells in the spleen without significantly changing the recombination pattern of *Tcf21*-Cre/ER in the spleen, bone marrow or liver. a–**v**, Pregnant female mice were at gestation day 18.5. **a**, **b**, H&E staining showed increased haematopoiesis in the spleens of pregnant mice (**b**) as evidenced by the presence of megakaryocytes (arrows; $n = 3$ mice per condition from 3 independent experiments). **c**–**i**, Pregnancy significantly increased spleen cellularity (**c**), as well as the numbers of HSCs (**d**), MEPs (**e**, **f**), Ter119$^+$ erythroid cells (**g**) and Gr-1$^+$Mac-1$^+$ myeloid cells (**h**) in the spleen but not the number of B220$^+$ or CD3$^+$ lymphoid cells (**i**). **j**, **k**, During pregnancy, *Scf*-GFP was expressed by VE-cadherin$^+$ endothelial cells and VE-cadherin$^-$ stromal cells (**j**) while *Cxcl12*-DsRed was expressed by a subset of the VE-cadherin$^-$*Scf*-GFP$^+$ stromal cells (**j**, **k**). **l**, Whole-mount imaging of a thick spleen section from a pregnant *Scf*^{GFP}; *Cxcl12*^{DsRed} mouse (representative images from 3 mice in 3 independent experiments). **m**, **n**, In the spleen, the numbers of *Scf*-GFP$^+$ cells (**m**) and *Cxcl12*-DsRed$^+$ cells (**n**) significantly increased upon bleeding. **o**, Endothelial and stromal cells in the spleen proliferated after bleeding. BrdU was administered to *Scf*^{GFP} mice or *Cxcl12*^{DsRed} mice for 18 days, beginning in pregnant mice after the plug was observed. The number of mice per treatment is indicated on each bar. Each panel shows mean ± s.d. from 3 independent experiments. Two-tailed Student's *t*-tests were used to assess statistical significance ($^{**}P < 0.01$, $^{***}P < 0.001$). **p**–**r**, Pregnancy did not alter the general distribution (**p**), perivascular localization (**q**) or surface marker expression (**r**; PDGFR-β$^+$ and LepR$^-$) of Tomato$^+$ cells in the spleens of *Tcf21*$^{cre/ER}$; *R26*^{tdTomato} mice. **s**, **t**, Tomato expression remained undetectable in the bone marrow of pregnant *Tcf21*$^{cre/ER}$; *R26*^{tdTomato} mice. **u**, **v**, During pregnancy, *Tcf21*-Cre/ER recombined in rare perivascular cells in the liver. **p**–**v**, $n = 3$ mice per genotype from 3 independent experiments.



Extended Data Figure 8 | Conditional deletion of *Cxcl12* from *Tcf21*$^+$ stromal cells impairs EMH in the spleens of pregnant mice without significantly affecting bone marrow haematopoiesis. a–x, Four-to-six-month-old female mice that had been treated with tamoxifen at least 2 months before were mated with normal wild-type males. Normal females and pregnant females at gestation day 18.5 were analysed. **a–d**, Conditional deletion of *Scf* from *Tcf21*$^+$ cells did not significantly affect the numbers of GMPs (**a**), CMPs (**b**), MEPs (**c**), Ter119$^+$ (erythroid), Gr-1$^+$Mac-1$^+$ (myeloid), CD3$^+$ (T) and B220$^+$ (B) cells (**d**) in one femur or one tibia. **e, f**, Conditional deletion of *Scf* from *Tcf21*$^+$ cells significantly reduced GMPs (**e**), Ter119$^+$ erythrocytes and Gr-1$^+$Mac-1$^+$ myeloid cells (**f**) in the spleen. **g, h**, Conditional deletion of *Scf* from *Tcf21*$^+$ cells did not significantly affect WBC (**g**) or platelet counts (**h**). **i–n**, Conditional deletion of *Cxcl12* from *Tcf21*$^+$ cells did not significantly affect bone marrow cellularity (**i**), or the numbers of HSCs (**j**), GMPs (**k**) CMPs (**l**), MEPs (**m**), Ter119$^+$ (erythroid), Gr-1$^+$Mac-1$^+$ (myeloid), CD3$^+$ (T) and B220$^+$ (B) cells (**n**) in the bone marrow. **o–w**, Spleen cellularity (**o**) and numbers of HSCs (**p**), GMPs (**q**), CMPs (**r**), MEPs (**s**), Ter119$^+$ (erythroid), Gr-1$^+$Mac-1$^+$ (myeloid), CD3$^+$ (T) and B220$^+$ (B) cells (**t**) in the spleen and WBC (**u**), platelet (**v**) and RBC counts (**w**) in the blood. **x**, Conditional deletion of *Cxcl12* from *Tcf21*$^+$ cells in the spleens of pregnant mothers did not significantly affect fetal mass. The numbers of mice per treatment are shown in each bar within each panel. Each panel shows mean ± s.d. from 3 independent experiments. **a–w**, The statistical significance of differences among genotypes was assessed using a repeated measures one-way ANOVA with Greenhouse–Geisser correction along with Tukey's multiple comparison tests with individual variances **x**, The statistical significance of differences was assessed using a two-way ANOVA. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, statistical significance relative to normal mice. $\dagger P < 0.05$, $\dagger\dagger P < 0.01$, $\dagger\dagger\dagger P < 0.001$, statistical significance between *Scf* mutant mice and control mice after bleeding.



Extended Data Figure 9 | Bleeding induces EMH and the proliferation of endothelial cells and stromal cells in the spleen without significantly changing the recombination pattern of *Tcf21*-Cre/ER in the spleen, bone marrow, or liver. a, **b**, H&E staining showed an increase in haematopoiesis in the spleen after repeated bleeding (**b**; bled) as evidenced by the presence of megakaryocytes (arrows; $n = 3$ mice per condition from 3 independent experiments). **c**–**i**, Bleeding significantly increased spleen cellularity (**c**), as well as the numbers of HSCs (**d**), MEPs (**e**, **f**), and the numbers of Ter119$^+$ erythroid cells (**g**) and Gr-1$^+$Mac-1$^+$ myeloid cells (**h**) in the spleen but not the number of B220$^+$ or CD3$^+$ lymphoid cells (**i**). **j**, **k**, After EMH induced by bleeding, *Scf*-GFP was expressed by VE-cadherin$^+$ endothelial cells and VE-cadherin$^-$ stromal cells (**j**) while *Cxcl12*-DsRed was expressed by a subset of the VE-cadherin$^-$*Scf*-GFP$^+$ stromal cells (**j**, **k**). **l**, Whole-mount imaging of a thick spleen section from a *Scf*^{GFP}; *Cxcl12*^{DsRed} mouse after bleeding (representative images from 3 mice in 3 independent experiments). **m**, **n**, The numbers of *Scf*-GFP$^+$ cells (**m**) and *Cxcl12*-DsRed$^+$ cells (**n**) significantly increased upon bleeding. **o**, Endothelial and stromal cells in the spleen proliferated after bleeding. BrdU was administered to *Scf*^{GFP} mice or *Cxcl12*^{DsRed} mice for 15 days beginning after the first bleeding. The numbers of mice per treatment are shown in each bar in each panel. Each panel shows mean ± s.d. from three independent experiments. Two-tailed Student's *t*-tests were used to assess statistical significance (**$P < 0.01$, ***$P < 0.001$). **p**–**r**, Bleeding did not alter the general distribution (**p**), perivascular localization (**q**) or surface marker expression (**r**; PDGFR-β$^+$ and LepR$^-$) of Tomato$^+$ cells in the spleens of *Tcf21*$^{cre/ER}$; *R26*^{tdTomato} mice. **s**, **t**, Tomato expression remained undetectable in bone marrow from *Tcf21*$^{cre/ER}$; *R26*^{tdTomato} mice after bleeding. **u**, **v**, After bleeding, *Tcf21*-Cre/ER recombined only in rare perivascular cells in the liver (**p**–**v**; $n = 3$ mice per genotype from 3 independent experiments).



Extended Data Figure 10 | Blood loss does not significantly change the recombination pattern of *Vav1*-Cre in the spleen, bone marrow, or liver; conditional deletion of *Cxcl12* from *Tcf21*$^+$ spleen stromal cells in bled mice impairs EMH in the spleen without significantly affecting bone marrow haematopoiesis. a–e′,Four-to-six-month-old mice with the indicated genetic backgrounds were repeatedly bled over a 2-week period. **a**–**h**, After EMH induced by blood loss, *Vav1*-Cre recombined efficiently in VE-cadherin$^+$ endothelial cells in the red pulp of the spleen (**a**–**c**) but poorly in the bone marrow (**d**–**f**) and liver (**g**, **h**), similar to what we observed under normal conditions (see Fig. 4a–c and Extended Data Fig. 5b) (**a**–**h**; $n = 3$ mice from 3 independent experiments). **i**–**n**, Conditional deletion of *Scf* using *Tcf21*-Cre/ER and/or *Vav1*-Cre did not significantly affect the numbers of GMPs (**i**), CMPs (**j**), MEPs (**k**), Ter119$^+$ (erythroid), Gr-1$^+$Mac-1$^+$ (myeloid), CD3$^+$ (T) and B220$^+$ (B) cells (**l**) in the bone marrow or WBC (**m**) or platelet counts in the blood (**n**). **o**, **p**, Conditional deletion of *Scf* using *Tcf21*-Cre/ER and/or *Vav1*-Cre significantly reduced GMPs (**o**), Ter119$^+$ erythrocytes and Gr-1$^+$Mac-1$^+$ myeloid cells (**p**) in the spleen. **i**–**p**, Data represent mean ± s.d. from 3 independent experiments. **q**–**v**, Conditional deletion of *Cxcl12* from *Tcf21*$^+$ spleen cells did not significantly affect bone marrow cellularity (**q**), or the numbers of HSCs (**r**), GMPs (**s**) CMPs (**t**), MEPs (**u**), Ter119$^+$ (erythroid), Gr-1$^+$Mac-1$^+$ (myeloid), CD3$^+$ (T) and B220$^+$ (B) cells (**v**) in one femur and one tibia from bled mice. **w**–**b′**, Conditional deletion of *Cxcl12* from *Tcf21*$^+$ spleen cells significantly reduced spleen cellularity (**w**), and the numbers of MEPs (**a′**) and erythroid cells (**b′**) in the spleens of bled mice. Conditional deletion of *Cxcl12* from *Tcf21*$^+$ spleen cells did not significantly affect the numbers of HSCs (**x**), GMPs (**y**), or CMPs (**z**) in the spleens of bled mice. **c′**–**e′**, Conditional deletion of *Cxcl12* from *Tcf21*$^+$ spleen cells significantly reduced RBC (**c′**) but not WBC (**d′**) or platelet counts (**e′**) in the blood of mice that had been repeatedly bled. **q**–**e′**, Data represent mean ± s.d. from 3 independent experiments. The numbers of mice per treatment are shown in each bar in each panel. Statistical significance of differences among genotypes was assessed using a repeated measures one-way ANOVA with Greenhouse–Geisser correction along with Tukey's multiple comparison tests with individual variances. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, statistical significance relative to normal mice. $\dagger P < 0.05$, $\dagger\dagger P < 0.01$, $\dagger\dagger\dagger P < 0.001$, statistical significance between *Scf* mutant mice and control mice after bleeding.

Extended Data Table 1 | Genes that are significantly more highly expressed by *Scf*-GFP$^+$ stromal cells in spleen as compared to bone marrow

Gene	Gene name	Unigene	Spleen *Scf*-GFP$^+$	BM *Scf*-GFP$^+$	Fold change
Coch	Coagulation factor C homolog	Mm.21325	12.1±0.3	6.6±0.0	45.4
Ccl21a	Chemokine (C-C motif) ligand 21A	Mm.458815	12.5±0.1	7.1±0.4	41.1
Acta2	Actin, alpha 2, smooth muscle, aorta	Mm.213025	11.9±0.3	6.7±0.1	35.2
Cxcl13	Chemokine (C-X-C motif) ligand 13	Mm.10116	11.8±0.3	6.8±0.2	30.3
Tcf21	Transcription factor 21	Mm.16497	11.3±0.6	6.6±0.0	25.9
Clca1	Chloride channel calcium activated 1	Mm.454553	11.1±0.3	6.6±0.0	22.5
Ifi27l2a	Interferon, alpha-inducible protein 27 like 2A	Mm.271275	11.3±0.2	7.2±0.4	16.6
Pln	Phospholamban	Mm.34145	10.7±0.1	6.6±0.0	16.3
Parm1	Prostate androgen-regulated mucin-like 1	Mm.5002	10.8±0.3	6.8±0.1	16
Fn1	Fibronectin 1	Mm.193099	10.7±0.4	6.8±0.2	14.9
Col14a1	Collagen, type XIV, alpha 1	Mm.297859	10.4±0.2	6.7±0.1	12.6
Nr4a1	Nuclear receptor subfamily 4, group A, 1	Mm.119	10.5±0.6	7.0±0.3	11.2
Agtr1a	Angiotensin II receptor, type 1a	Mm.35062	10.7±0.7	7.3±0.6	11
Fos	FBJ osteosarcoma oncogene	Mm.246513	11.4±0.4	8.0±0.4	10.7
Atp1b2	ATPase, Na+/K+ transporter, beta 2	Mm.235204	10.6±0.2	7.2±0.2	10.6
Tnxb	Tenascin XB	Mm.290527	9.9±0.5	6.6±0.0	9.5
Myh11	Myosin, heavy polypeptide 11, smooth muscle	Mm.250705	10.7±0.7	7.5±0.2	9.4
Hspb1	Heat shock protein 1	Mm.13849	10.8±0.7	7.6±0.2	9.3
Clca2	Chloride channel calcium activated 2	Mm.20897	9.8±0.4	6.6±0.0	8.8
Tagln	Transgelin	Mm.283283	10.4±0.5	7.3±0.9	8.6
Nr2f2	Nuclear receptor subfamily 2, group F, 2	Mm.158143	10.7±0.3	7.6±0.3	8.5
Mustn1	Musculoskeletal, embryonic nuclear protein 1	Mm.220895	10.8±0.5	7.7±0.7	8.2
Aspn	Asporin	Mm.383216	9.7±0.6	6.6±0.0	8.2
Sparcl1	SPARC-like 1	Mm.29027	12.1±0.1	9.1±0.4	8.1

Significance was considered as >8 fold and $P < 0.015$. Data show mean ± s.d. for $\log_2$ transformed expression values ($n = 3$ independent samples per cell population). Maximal background expression was considered to be 6.6 ($\log_2(100)$); all expression values below this threshold were set to 6.6 for purposes of calculating fold change. Two-tailed Student's *t*-tests were used to assess statistical significance. Data for bone marrow *Scf*-GFP$^+$ stromal cells are from ref. 19.

ARTICLE

doi:10.1038/nature15748

Epithelial–to–mesenchymal transition is not required for lung metastasis but contributes to chemoresistance

Kari R. Fischer[1,2,3,4], Anna Durrans[1,2,3], Sharrell Lee[1,2,3], Jianting Sheng[5], Fuhai Li[5], Stephen T. C. Wong[5,6], Hyejin Choi[1,2,3,4], Tina El Rayes[1,2,3,4], Seongho Ryu[1,3,7], Juliane Troeger[8,9], Robert F. Schwabe[8,9], Linda T. Vahdat[10], Nasser K. Altorki[1,3], Vivek Mittal[1,2,3] & Dingcheng Gao[1,2,3]

The role of epithelial–to–mesenchymal transition (EMT) in metastasis is a longstanding source of debate, largely owing to an inability to monitor transient and reversible EMT phenotypes *in vivo*. Here we establish an EMT lineage–tracing system to monitor this process in mice, using a mesenchymal–specific Cre–mediated fluorescent marker switch system in spontaneous breast–to–lung metastasis models. We show that within a predominantly epithelial primary tumour, a small proportion of tumour cells undergo EMT. Notably, lung metastases mainly consist of non–EMT tumour cells that maintain their epithelial phenotype. Inhibiting EMT by overexpressing the microRNA miR–200 does not affect lung metastasis development. However, EMT cells significantly contribute to recurrent lung metastasis formation after chemotherapy. These cells survived cyclophosphamide treatment owing to reduced proliferation, apoptotic tolerance and increased expression of chemoresistance–related genes. Overexpression of miR–200 abrogated this resistance. This study suggests the potential of an EMT–targeting strategy, in conjunction with conventional chemotherapies, for breast cancer treatment.

Despite significant advances in diagnosing and treating cancer, metastasis persists as a barrier to successful therapy and the main cause of cancer-related death[1]. The EMT, wherein epithelial cells depolarize, lose their cell–cell contacts, and gain an elongated, fibroblast-like morphology, is a potential mechanism by which tumour cells gain metastatic features. Functional implications of EMT include enhanced mobility, invasion and resistance to apoptotic stimuli[2,3]. Moreover, through EMT tumour cells acquire cancer stem cell, secondary tumour-initiating and chemoresistance properties[4–6]. However, the importance of EMT *in vivo* is fiercely debated owing to major challenges. Mesenchymal tumour cells cannot easily be distinguished from neighbouring stromal cells, and metastatic lesions mostly exhibit epithelial phenotypes[7]. The latter may be due to the hypothesized reverse process, mesenchymal to epithelial transition (MET), of the disseminated tumour cells. Studies have confirmed that mesenchymal cells are more capable of escaping the primary tumour, and of reaching distant sites, but it remains unproven that those same cells complete the full metastatic cascade in the form of a secondary nodule. Without evidence for the dissemination, colonization and metastatic outgrowth of mesenchymal tumour cells, the role of EMT will remain contested. In this study, we employed multiple transgenic mouse models, establishing a cell lineage tracing approach together with characterization of epithelial and mesenchymal markers, to address the requirement of EMT in metastasis. The newly established transgenic model also provided us a unique opportunity to study the contribution of EMT to chemoresistance.

EMT lineage tracing during metastasis

To track EMT during metastasis *in vivo*, we generated a mesenchymal-specific, Cre-mediated fluorescent marker switch strategy and established a triple-transgenic mouse model (*MMTV-PyMT/Rosa26-RFP-GFP/Fsp1-cre*, tri-PyMT, Fig. 1a). In these mice, spontaneous multifocal breast adenocarcinomas with distinct epithelial characteristics resembling the human luminal subtype develop in the mammary glands, and give rise to lung metastases with high penetrance[8,9]. The *Fsp1* (fibroblast specific protein 1) promoter drives expression of Cre recombinase in cells of mesenchymal lineage[10]. A Cre-switchable fluorescent marker (*lox*-RFP-STOP-*lox*-GFP) is ubiquitously expressed under the control of the β-actin promoter in the *Rosa26* locus[11]. *Fsp1* is the critical gatekeeping gene of EMT initiation[12], and its early activation in this process[13] allows for lineage tracing of tumour cells that have undergone EMT *in vivo*. Importantly, the colour switch system is irreversible—even if the mesenchymal tumour cells undergo MET in the metastatic organs[14], they would remain GFP^+.

Primary breast tumours developed in the tri-PyMT mice at 8 weeks of age. Immunofluorescence revealed that the majority of tumour cells, identified by PyMT oncogene expression, were RFP positive (Fig. 1b). These cells expressed E-cadherin and lacked vimentin (Extended Data Fig. 1a), indicating their epithelial phenotype. The GFP^+ cells detected in the tumour bed were largely haematopoietic cells as they are PyMT negative and express CD45, a pan-haematopoietic marker (Extended Data Fig. 1a), which is consistent with previous reports[15]. Altogether,

[1]Department of Cardiothoracic Surgery, Weill Cornell Medical College of Cornell University, 1300 York Avenue, New York, New York 10065, USA. [2]Department of Cell and Developmental Biology, Weill Cornell Medical College of Cornell University, 1300 York Avenue, New York, New York 10065, USA. [3]Neuberger Berman Lung Cancer Center, Weill Cornell Medical College of Cornell University, 1300 York Avenue, New York, New York 10065, USA. [4]Weill Cornell Graduate School of Medical Sciences, Weill Cornell Medical College of Cornell University, 1300 York Avenue, New York, New York 10065, USA. [5]Department of Systems Medicine and Bioengineering, Houston Methodist Research Institute, Houston Methodist Hospital, Houston, Texas 77030, USA. [6]Methodist Cancer Center, Houston Methodist Hospital, Houston, Texas, 77030 USA. [7]Soonchunhyang Institute of Medi-bio Science (SIMS), Soonchunhyang University, 25 Bongjeong-ro Cheonan-Si, Chungcheongnam-do 31151, South Korea. [8]Department of Medicine, Columbia University, College of Physicians and Surgeons, New York, New York 10032, USA. [9]Institute of Human Nutrition, Columbia University, New York, New York 10032, USA. [10]Department of Medicine, Weill Cornell Medical College of Cornell University, 1300 York Avenue, New York, New York 10065, USA.

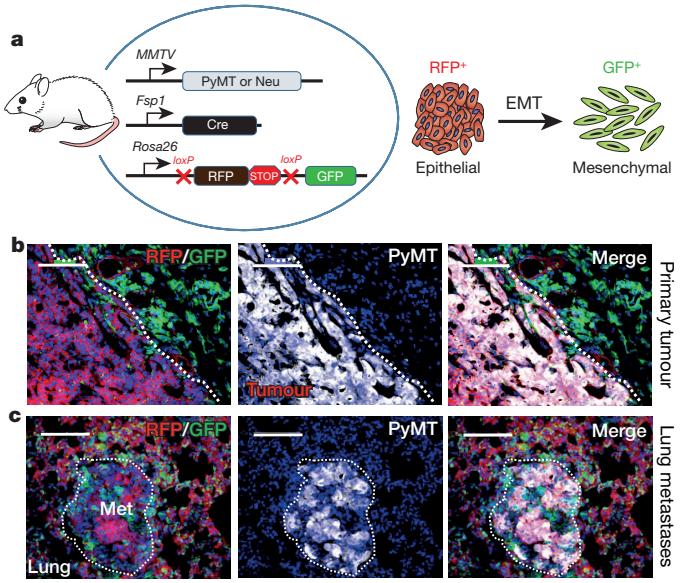


Figure 1 | Establishing an EMT lineage tracing system in triple-transgenic mice. a, Schematic of triple-transgenic mice carrying polyoma middle-T (PyMT) or Neu oncogenes driven by the *MMTV* promoter, Cre recombinase under the control of the *Fsp1* promoter, and floxed RFP-STOP followed by GFP under control of the β-actin promoter in the *Rosa26* locus. RFP^+ epithelial tumour cells undergoing EMT permanently convert into GFP^+ cells following activation of *Fsp1*–Cre. **b**, **c**, Immunofluorescent microscopy images of tri-PyMT primary tumours (**b**) and lung metastases (met; **c**) (>10 sections from 3 mice), depicting RFP^+ and GFP^+ cells within the tumour bed, and staining (white, pseudo-coloured) for PyMT. Scale bars, 100 μm.

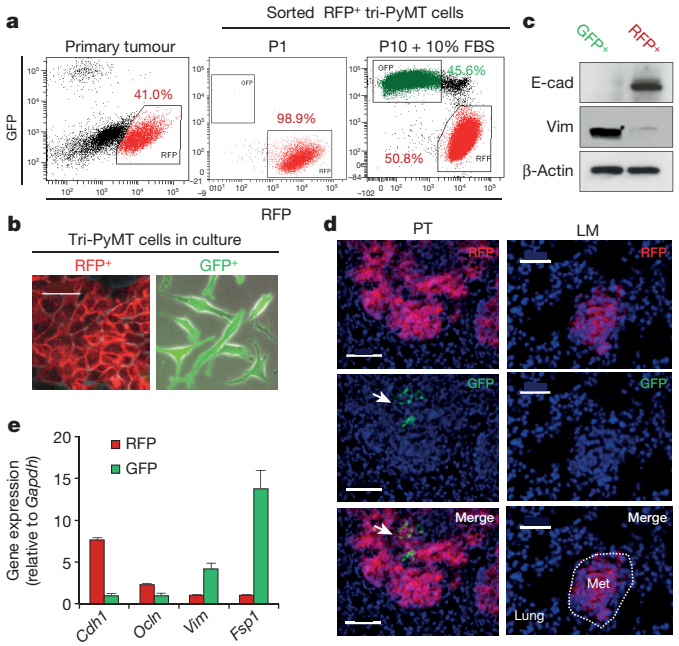


Figure 2 | The EMT lineage tracing system reports EMT in tumour cells with high fidelity. a, Scatter plots from flow cytometry analysis of tri-PyMT primary tumour cells, depicting GFP^+ and RFP^+ populations in the primary tumour immediately after sorting of RFP^+ cells (P1), and after ten passages in culture with 10% FBS (P10 + 10% FBS). Numbers indicate the percentage of RFP^+ and GFP^+ cells in the total population. **b**, Phase contrast/fluorescent overlay image of tri-PyMT cells in culture. Scale bar, 50 μm. **c**, Western blot of sorted RFP^+ and GFP^+ tri-PyMT cells for E-cadherin, vimentin and β-actin as a loading control. Representative of two individual experiments. For original gel images, see Supplementary Fig. 1. **d**, Representative imaging of GFP^+ and RFP^+ tumour cells in primary tumours (PT) and lung metastases (LM) in the orthotopic model ($n = 8$ mice). Arrow indicates scattered GFP^+ EMT tumour cells in the primary tumour. Scale bars, 100 μm (PT) and 50 μm (LM). **e**, qRT–PCR analysis of relative expression of EMT markers in RFP^+ and GFP^+ cells sorted from orthotopic tri-PyMT primary tumours. *Gapdh* served as the internal control. E-cadherin is encoded by the *Cdh1* gene. Occludin is encoded by the *Ocln* gene. Data are reported as mean ± s.e.m., $n = 4$ primary tumours.

this data suggests that tumour cells maintain their original RFP expression and epithelial phenotype in the primary tumour.

Lung metastasis developed spontaneously in tri-PyMT lungs at 12 weeks of age. Surprisingly, the PyMT-positive metastatic lesions were RFP^+ (Fig. 1c), and epithelial ($E\text{-cadherin}^+/\text{vimentin}^-$) (Extended Data Fig. 1b), whereas only non-tumour cells expressed GFP. These results indicate that tumour cells did not activate the mesenchymal-specific *Fsp1* promoter, and retained their epithelial phenotype during metastasis. Thus, tumour cells may not undergo EMT to form metastatic lesions.

Lineage tracing in additional models

To exclude the possibility that the absence of EMT in metastasis may be unique to PyMT-driven breast tumours, we established EMT lineage tracing in the Neu oncogene-driven[16] spontaneous breast cancer model (*MMTV-neu/Rosa26-RFP-GFP/Fsp1-Cre*, tri-Neu mouse). The Neu (ErbB-2) proto-oncogene is associated with 20–30% of human breast cancers, and *MMTV*–Neu transgenic mice spontaneously develop focal adenocarcinomas resembling human luminal phenotypes after an extended latency at 6–8 months of age. Lung metastases are frequently (72%) observed in these transgenic mice at 9–12 months of age. Mirroring the tri-PyMT model, the Neu^+ tumour cells in both primary and metastatic lesions in tri-Neu mice were also RFP^+ and epithelial ($E\text{-cad}^+/\text{Vim}^-$) (Extended Data Fig. 2). Therefore, the absence of EMT during metastasis formation is an oncogene-independent phenomenon, manifesting in both PyMT and Neu-driven tumours.

To overcome the limitation of using solely *Fsp1*–Cre to indicate EMT, we acquired the vimentin–CreER transgenic mouse, which successfully traced mesenchymal lineage cells during liver fibrosis[17], and generated an additional EMT lineage tracing model (tri-PyMT/Vim mice, *MMTV-PyMT/Rosa26-RFP-GFP/Vimentin-creER*). After continuous induction of Cre activity by Tamoxifen injection (2 mg, intraperitoneal, three times per week starting when the primary tumours appear at 8 weeks of age) the majority of tumour cells in both the primary and metastatic lesions in tri-PyMT/Vim mice were RFP^+ (Extended Data Fig. 3)—suggesting an absence of vimentin promoter activation during lung metastasis formation. EMT marker staining also revealed the epithelial phenotype ($E\text{-cad}^+/\text{Vim}^-$) of the tumour cells in both primary and metastatic lesions (Extended Data Fig. 3).

Together, results from two oncogene-driven metastatic tumour models (*MMTV*–PyMT and *MMTV*–Neu) and two independent mesenchymal-specific reporters (*Vim*–Cre and *Fsp1*–Cre) suggest that EMT does not significantly contribute to the development of lung metastases.

Validating EMT lineage tracing

To evaluate the specificity and sensitivity of the EMT lineage tracing system, we established a cell line from the tri-PyMT breast tumours. In culture, RFP^+ tri-PyMT cells switched their fluorescent marker expression to GFP, as indicated by the presence of a RFP^+/GFP^+ double-positive transitioning population (Fig. 2a). The cells were cultured in 10% FBS, and serum is known to be enriched for many EMT promoting factors including TGFs[18]. Moreover, addition of TGF-β1 in low-serum conditions (2% FBS), yielded an increase in GFP^+ cells (Extended Data Fig. 4a). In concert with the fluorescent marker switch, tri-PyMT cells changed their morphology from cobblestone-like clusters of epithelial cells to dispersed spindle-shaped mesenchymal cells (Fig. 2b). Reflecting the morphologic differences, the GFP^+ cells were more motile than RFP^+ cells (Extended Data Fig. 4b).

The fidelity of the EMT lineage tracing system was confirmed by analysis of EMT marker expression in sorted RFP^+ and GFP^+ tri-PyMT cells. RFP^+ cells expressed elevated levels of epithelial markers including E-cadherin and Occludin, while GFP^+ cells expressed

several mesenchymal markers including vimentin, FSP1, Twist, Zeb1 and Zeb2 as determined by quantitative reverse transcription PCR (qRT–PCR) (Extended Data Fig. 4c). Both RFP$^+$ and GFP$^+$ tri-PyMT cells expressed the PyMT oncogene. Consistently, western blot analysis confirmed the differential expression of E-cadherin and vimentin in RFP$^+$ and GFP$^+$ cells (Fig. 2c). Flow cytometry for E-cadherin revealed that the majority of E-cadherin$^-$ cells were GFP$^+$ (97.4%) (Extended Data Fig. 4d). Of note, the E-cadherin$^+$ cells were either RFP$^+$ (93.6%) or RFP$^+$/GFP$^+$ (6.0%), demonstrating that tumour cells switch their fluorescent marker expression before the loss of epithelial markers, and validating the early reporting of EMT in our system. These results confirm that the *Fsp1*–Cre-mediated fluorescent marker switch in tumour cells reports EMT with high fidelity and efficiency.

Rare EMT events in tumour progression

In the triple-transgenic models, ubiquitous expression of GFP in the tumour microevironment precluded detection of potentially rare GFP$^+$ tumour cells. To confine the fluorescence to tumour cells, we established an orthotopic model by implanting purified RFP$^+$ tri-PyMT cells in wild-type mice (Extended Data Fig. 5a, b). Consistent with observations in the triple-transgenic mice, primary tumours contained RFP$^+$ epithelial cells (Fig. 2d). However, GFP$^+$ cells were detected, indicating tumour EMT (Fig. 2d and Extended Data Fig. 5c, upper panel). These cells lacked E-cadherin (Extended Data Fig. 5c, upper panel) and made up 1.98 ± 1.40% ($n = 6$) of the total tumour cells (Extended Data Fig. 5d). qRT–PCR analysis of EMT markers comparing sorted RFP$^+$ and GFP$^+$ cells from the same primary tumour confirmed the mesenchymal phenotype of the GFP$^+$ cells (Fig. 2e). Importantly, these GFP$^+$ EMT tumour cells did not contribute to lung metastasis. Early disseminated tumour cells detected in the lungs were epithelial and RFP$^+$ (Extended Data Fig. 5c, middle panel), and 28 lung nodules detected in 8 mice maintained the epithelial phenotype (Fig. 2d and Extended Data Fig. 5c, lower panel).

We also established an orthotopic tri-PyMT/Vim model, wherein Tamoxifen was administered directly after orthotopic injection to ensure immediate tracing of EMT events. Consistently, the majority of tumour cells in both primary and metastatic tumours were RFP$^+$ and epithelial (Extended Data Fig. 6). Again, GFP$^+$ EMT events (4.46 ± 1.0% of total tumour cells, $n = 3$) were detected in the primary tumours.

To further dissect the metastatic cascade, we quantified the relative numbers of RFP$^+$ and GFP$^+$ cells in the primary tumour, blood and metastases of the tri-PyMT orthotopic model by flow cytometry. An RFP to GFP ratio of ~100:1 in the primary tumour and ~15:1 in the blood was observed (Extended Data Fig. 7a, b). However, gain by the enrichment of GFP$^+$ cells in circulation did not translate to an advantage in metastatic outgrowth, as the RFP:GFP ratio in the lung was ~150:1. Altogether, these findings are consistent with our observations in the triple-transgenic models, suggesting that the majority of breast tumour cells persist in an epithelial state during primary tumour growth and lung metastasis formation.

EMT inhibition and metastasis formation

In spite of the extensive characterization of the EMT reporter system, there was still the distant possibility of our reporter failing to manifest all EMT events *in vivo*. Therefore, we sought to inhibit EMT and determine its impact on metastasis. We ectopically expressed miR-200, a well-known inhibitor of EMT that directly targets *Zeb1* and *Zeb2*—the transcriptional repressors of E-cadherin[19,20]. We posited that stably expressing miR-200 in tri-PyMT cells would block EMT and trap tumour cells in a permanent epithelial state. Compared with control cells, miR-200 overexpressing cells (Extended Data Fig. 7c) showed elevated expression of epithelial cell markers and reduced expression of mesenchymal markers (Extended Data Fig. 7d). As expected, overexpression of miR-200 inhibited the RFP to GFP conversion (>90% remaining RFP$^+$, Fig. 3a). These results substantiate effective miR-200 suppression of EMT in the tri-PyMT cells.

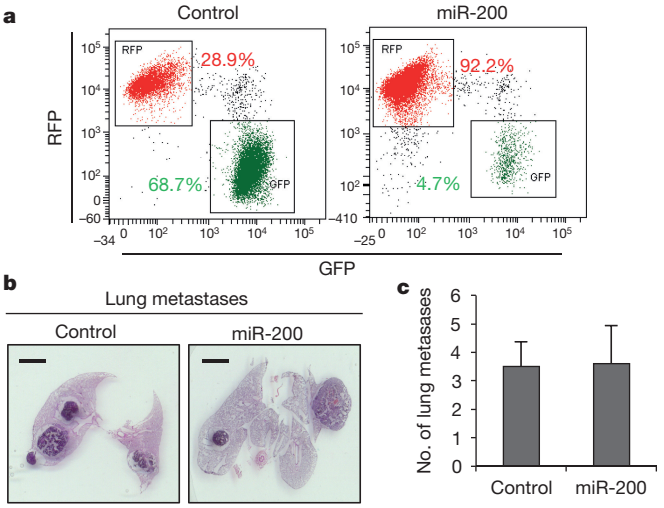


Figure 3 | mir-200 inhibition of EMT in tri-PyMT cells did not impact lung metastasis. a, Flow cytometry analysis of tri-PyMT control and mir-200-expressing cells, indicating the percentage of RFP$^+$ and GFP$^+$ cells. **b**, Representative histologic lung images in tri-PyMT control and mir-200-expressing orthotopic mice ($n = 5$). Scale bar, 1.5 mm. **c**, Quantification of lung metastasis formation (number of individual nodules) in tri-PyMT control and mir-200-expressing tumour-bearing mice ($n = 5$). Data reported as the mean ± s.e.m.

To explore the impact of inhibiting EMT on metastasis formation *in vivo*, we orthotopically injected miR-200 overexpressing tri-PyMT cells. We identified 18 metastases in 5 mice, a similar ratio to that observed in mice bearing control tri-PyMT cells (28 metastases in 8 mice) (Fig. 3b, c). These results demonstrate that inhibition of EMT by miR-200 overexpression does not impair the ability of tumour cells to form distant lung metastases.

EMT is involved in chemoresistance

Emerging evidence suggests a molecular and phenotypic association between EMT and chemoresistance in several cancers[21–23]. Compellingly, residual breast cancers following chemotherapy display a mesenchymal phenotype and tumour-initiating features[23]. To determine if the acquisition of chemoresistance induces specific molecular changes consistent with EMT, we evaluated the orthotopic tri-PyMT model under chemotherapy. Animals with established primary tumours were treated with cyclophosphamide (CTX), a commonly used drug in breast cancer treatment[24] (100 mg kg^{-1}, once per week, for two weeks prior, and two weeks after, surgery; Fig. 4a). The tumours responded to chemotherapy, manifesting a 60% reduction in growth and markedly enhanced apoptotic activity (Extended Data Fig. 8a–c). Of note, the RFP$^+$ cells were highly proliferative and apoptotic in comparison with GFP$^+$ cells in CTX-treated mice (Extended Data Fig. 8d–g), suggesting that GFP$^+$ cells have reduced susceptibility to chemotherapy. However, in the primary tumour, the GFP$^+$ cell percentage remained static under CTX treatment (Extended Data Fig. 8h).

Remarkably, in the early metastatic lungs (four weeks after tumour inoculation), flow cytometry analysis revealed a 2.7:1 ratio of GFP$^+$ to RFP$^+$ cells in CTX-treated mice (Fig. 4b). Subsequently at four weeks after cessation of treatment, a notable contribution of GFP$^+$ tumour cells was detected in 5 out of 17 metastatic lesions (Fig. 4a). This is in contrast to untreated mice, where all metastatic lesions were derived from RFP$^+$ cells (Fig. 2d), suggesting that the EMT process may be involved in metastatic outgrowth in the context of chemotherapy.

To evaluate the effects of CTX on the EMT and non-EMT cell populations, sorted GFP$^+$ and RFP$^+$ cells were incubated with CTX *in vitro*—the GFP$^+$ cells were markedly more resistant to both short- and long-term treatment (Fig. 4c and Extended Data Fig. 9a, b). The selective advantage of mesenchymal tumour cells in the context of chemotherapy

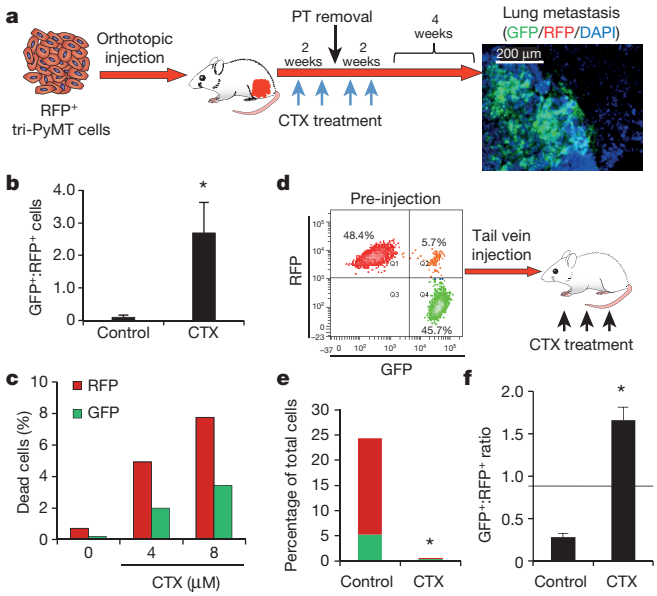


Figure 4 | EMT tumour cells are resistant to chemotherapy. a, Schema of CTX treatment in tri-PyMT orthotopic model. Mice bearing an RFP$^+$ primary tumour were treated with CTX (100 mg kg^{-1}, once per week, for 4 weeks, as indicated by blue arrows). After 2 weeks of treatment, primary tumour (PT) was removed (black arrow). Lung metastasis growth was permitted for 4 weeks post CTX treatment. Fluorescent imaging of lungs revealed the contribution of GFP$^+$ tumour cells to lung metastases ($n = 9$ mice). **b**, Ratio of GFP$^+$ to RFP$^+$ cells in early metastatic lungs (4 weeks post orthotopic injection) of untreated control and CTX-treated mice as quantified by flow cytometry ($n = 4$, $^*P < 0.05$). Data reported as the mean ± s.e.m. **c**, Apoptosis (as measured by Annexin binding) of RFP$^+$ and GFP$^+$ tri-PyMT cells treated with CTX ($n = 2$ biological replicates). **d**, Flow cytometry scatter plot showing the proportions of RFP$^+$ and GFP$^+$ tri-PyMT cells before intravenous injection. Mice were treated with CTX (100 mg kg^{-1} per week for 3 weeks, $n = 5$ mice per group). **e**, Quantification of flow cytometry data showing the percentage of RFP$^+$ and GFP$^+$ tumour cells (red and green bars, respectively) of total cells in the lung of control and CTX-treated mice ($n = 5$ mice per group, $^*P < 0.05$). **f**, Quantification of flow cytometry data showing the ratio of GFP$^+$ to RFP$^+$ cells in lungs of control and CTX-treated mice. Black line represents the starting ratio of GFP$^+$ to RFP$^+$ cells before injection as derived from the data in Fig. 4d ($^*P < 0.05$). Data reported as the mean ± s.e.m.

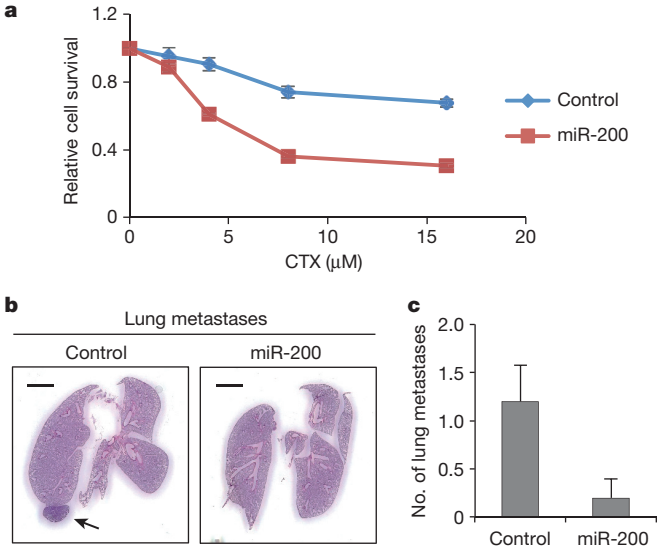


Figure 5 | miR-200 overexpression abrogates CTX resistance. **a**, Sensitivity of Control and miR-200-expressing tri-PyMT tumour cells to CTX treatment as measured by CellTiter-Glo. $n = 4$ biological replicates per condition **b**, Representative histologic lung images in tri-PyMT control and mir-200-expressing tumour-bearing mice treated with CTX ($n = 5$). Scale bar, 1.5 mm. **c**, Quantification of lung metastasis formation (number of individual nodules) in CTX-treated tri-PyMT control and mir-200-expressing tumour-bearing mice ($n = 5$). Data reported as the mean ± s.e.m.

was then corroborated by a competitive survival assay *in vivo* (Fig. 4d). Mice were injected intravenously with an equivalent number of RFP$^+$ and GFP$^+$ cells, and immediately received CTX (100 mg kg^{-1}, once per week). After three weeks, lungs were harvested and the ratio of RFP$^+$ and GFP$^+$ cells was assessed by flow cytometry. CTX significantly inhibited outgrowth of lung metastasis from both RFP$^+$ and GFP$^+$ cells (Fig. 4e). The untreated lungs were morbidly overwhelmed with tumours, with nearly 80% of the tumour cells detected as RFP$^+$. Conversely, in CTX-treated mice, more than 60% of the surviving tumour cells were GFP$^+$, producing a significantly higher ratio of GFP:RFP cells in these mice (Fig. 4f). These results indicate that GFP$^+$ EMT cells are more resistant to chemotherapy both *in vitro* and *in vivo*.

Immunostaining revealed that in the untreated mice, both RFP$^+$ and GFP$^+$ cells formed epithelial metastatic lesions (E-cad$^+$/Vim$^-$) (Extended Data Fig. 9c). Given the initial mesenchymal phenotypes of GFP$^+$ cells before injection, this suggests that the GFP$^+$ tumour cells have undergone MET in the metastatic organ. On the other hand, in CTX-treated mice the majority of surviving tumour cells were scattered mesenchymal GFP$^+$ cells (E-cad$^-$/Vim$^+$) (Extended Data Fig. 9d). Together, these observations suggest that EMT tumour cells that sustain a mesenchymal phenotype are resistant to chemotherapy.

To begin to investigate the molecular underpinnings of mesenchymal tumour cell resistance, we analysed the transcriptomic changes of EMT tumour cells. We sorted RFP$^+$ and GFP$^+$ cells and performed RNA-sequencing analysis (Supplementary Information Table 1). In addition to the expected changes in EMT marker expression (Extended Data Fig. 10a), the expression of many cell-proliferation-related genes was reduced in GFP$^+$ cells (Extended Data Fig. 10b), mirroring their phenotype of reduced proliferation *in vivo*. The GFP$^+$ cells also showed increased expression of proven chemoresistance-related factors including IL6, Periostin, Enpp2 and Pdgfr[25–28]. Additionally, the CTX-treated GFP$^+$ cells elevated their expression of many drug-metabolizing enzymes including drug transporters (Abcb1a, Abcb1b and Abcc1), aldehyde dehydrogenases (ALDHs), cytochrome P450s, and glutathione-metabolism-related enzymes (Extended Data Fig. 10c). The main toxicity of CTX is due to its metabolite phosphoramide mustard, which is only formed in cells with low levels of ALDHs. ALDH converts the CTX-metabolite aldophosphamide into the non-toxic carboxyphosphamide[29]. In accordance with the transcriptomic data, GFP$^+$ cells had significantly higher ALDH activity compared with RFP$^+$ cells (Extended Data Fig. 10d). These properties of reduced proliferation, increased apoptotic resistance, and upregulation of chemoresistance and drug metabolizing genes in GFP$^+$ EMT tumour cells may contribute to their insensitivity to CTX. Notably, GFP$^+$ cells were also refractory to other commonly used chemotherapies including doxorubicin, paclitaxel, and fluorouracil treatment (Extended Data Fig. 10e).

To demonstrate that the EMT is required for the generation of CTX resistance, we first tested *in vitro* the effect of treatment on control and miR-200 overexpressing tri-PyMT cells. With increasing concentrations of CTX, the miR-200 cells were significantly more susceptible to therapy (Fig. 5a). We then expanded upon this finding *in vivo*, establishing orthotopic control and miR-200 primary tumours, and applying the pre- and post-surgery CTX regimen. We found that by blocking EMT in tumour cells, we effectively ablated metastatic growth (Fig. 5b, c). Thus, EMT contributes to the development of chemoresistant metastasis.

Discussion

Using two independent EMT lineage tracing strategies in two disparate oncogene-driven autochthonous models of breast cancer, we demonstrated that lung metastases are derived from non-EMT tumour

cells, contradicting the original EMT/MET hypothesis[2,30]. In a tracing system similar to our own, EMT was identified in primary tumours, but the mesenchymal lineage status of the metastatic nodules was not pursued[31]. Ultimately in our models we found that tumour cells disseminate and form metastases while persisting in their epithelial phenotype, in accordance with a recent study[32]. To underline that EMT is not required for metastasis, overexpression of miR-200—a microRNA that is incongruously associated with both reduced invasion[19,20] and increased metastasis[33]—resulted in combined suppression of the EMT-promoting transcription factors Snail1/2, Twist, Zeb1 and Zeb2, but had no effect on metastasis. Given that both epithelial and mesenchymal tumour cells have the potential to disseminate, it is plausible that the larger fraction of highly proliferative epithelial cells outcompete the minor EMT tumour cell population in generating macrometastatic lesions.

Until now, the majority of data connecting EMT with chemoresistance was largely derived from *in vitro* studies, or clinical prognostic data. Here we demonstrate that highly proliferative non-EMT cells are sensitive to chemotherapy, and observe the emergence of recurrent EMT-derived metastases after treatment. There is a great emphasis towards developing EMT-targeting therapies[34,35], and our studies suggest that while EMT blockade may not affect metastasis formation, specifically targeting EMT tumour cells will be synergistic with conventional chemotherapy. Thus, our EMT lineage tracing system provides a unique preclinical platform to develop combination therapies that will eliminate both populations, and combat chemoresistance.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 17 August 2014; accepted 23 September 2015.
Published online 11 November 2015.

1. Bastid, J. EMT in carcinoma progression and dissemination: facts, unanswered questions, and clinical considerations. *Cancer Metastasis Rev.* **31,** 277–283 (2012).
2. Kalluri, R. & Weinberg, R. A. The basics of epithelial–mesenchymal transition. *J. Clin. Invest.* **119,** 1420–1428 (2009).
3. Scheel, C. & Weinberg, R. A. Phenotypic plasticity and epithelial–mesenchymal transitions in cancer and normal stem cells? *Int. J. Cancer* **129,** 2310–2314 (2011).
4. Mani, S. A. *et al.* The epithelial–mesenchymal transition generates cells with properties of stem cells. *Cell* **133,** 704–715 (2008).
5. Gal, A. *et al.* Sustained TGF beta exposure suppresses Smad and non-Smad signalling in mammary epithelial cells, leading to EMT and inhibition of growth arrest and apoptosis. *Oncogene* **27,** 1218–1230 (2008).
6. Hennessy, B. T. *et al.* Characterization of a naturally occurring breast cancer subset enriched in epithelial-to-mesenchymal transition and stem cell characteristics. *Cancer Res.* **69,** 4116–4124 (2009).
7. Gao, D. *et al.* Myeloid progenitor cells in the premetastatic lung promote metastases by inducing mesenchymal to epithelial transition. *Cancer Res.* **72,** 1384–1394 (2012).
8. Lin, E. Y. *et al.* Progression to malignancy in the polyoma middle T oncoprotein mouse breast cancer model provides a reliable model for human diseases. *Am. J. Pathol.* **163,** 2113–2126 (2003).
9. Guy, C. T., Cardiff, R. D. & Muller, W. J. Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease. *Mol. Cell. Biol.* **12,** 954–961 (1992).
10. Bhowmick, N. A. *et al.* TGF-β signaling in fibroblasts modulates the oncogenic potential of adjacent epithelia. *Science* **303,** 848–851 (2004).
11. Muzumdar, M. D., Tasic, B., Miyamichi, K., Li, L. & Luo, L. A global double-fluorescent Cre reporter mouse. *Genesis* **45,** 593–605 (2007).
12. Xue, C., Plieth, D., Venkov, C., Xu, C. & Neilson, E. G. The gatekeeper effect of epithelial–mesenchymal transition regulates the frequency of breast cancer metastasis. *Cancer Res.* **63,** 3386–3394 (2003).
13. Okada, H., Danoff, T. M., Kalluri, R. & Neilson, E. G. Early role of Fsp1 in epithelial–mesenchymal transformation. *Am. J. Physiol.* **273,** F563–F574 (1997).
14. Gunasinghe, N. P., Wells, A., Thompson, E. W. & Hugo, H. J. Mesenchymal–epithelial transition (MET) as a mechanism for metastatic colonisation in breast cancer. *Cancer Metastasis Rev.* **31,** 469–478 (2012).
15. Cabeżon, T. *et al.* Expression of S100A4 by a variety of cell types present in the tumor microenvironment of human breast cancer. *Int. J. Cancer* **121,** 1433–1444 (2007).
16. Guy, C. T. *et al.* Expression of the neu protooncogene in the mammary epithelium of transgenic mice induces metastatic disease. *Proc. Natl Acad. Sci. USA* **89,** 10578–10582 (1992).
17. Troeger, J. S. *et al.* Deactivation of hepatic stellate cells during liver fibrosis resolution in mice. *Gastroenterology* **143,** 1073–1083 (2012).
18. Dumont, N. *et al.* Sustained induction of epithelial to mesenchymal transition activates DNA methylation of genes silenced in basal-like breast cancers. *Proc. Natl Acad. Sci. USA* **105,** 14867–14872 (2008).
19. Park, S. M., Gaur, A. B., Lengyel, E. & Peter, M. E. The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. *Genes Dev.* **22,** 894–907 (2008).
20. Gregory, P. A. *et al.* The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nature Cell Biol.* **10,** 593–601 (2008).
21. Singh, A. & Settleman, J. EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer. *Oncogene* **29,** 4741–4751 (2010).
22. Zhang, Y., Toy, K. A. & Kleer, C. G. Metaplastic breast carcinomas are enriched in markers of tumor-initiating cells and epithelial to mesenchymal transition. *Mod. Pathol.* **25,** 178–184 (2012).
23. Creighton, C. J. *et al.* Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features. *Proc. Natl Acad. Sci. USA* **106,** 13820–13825 (2009).
24. von Minckwitz, G. Docetaxel/anthracycline combinations for breast cancer treatment. *Expert Opin. Pharmacother.* **8,** 485–495 (2007).
25. Lau, C. K. *et al.* An Akt/hypoxia-inducible factor-1α/platelet-derived growth factor-BB autocrine loop mediates hypoxia-induced chemoresistance in liver cancer cells and tumorigenic hepatic progenitor cells. *Clin. Cancer Res.* **15,** 3462–3471 (2009).
26. Xiao, Z. M., Wang, X. Y. & Wang, A. M. Periostin induces chemoresistance in colon cancer cells through activation of the PI3K/Akt/survivin pathway. *Biotechnol. Appl. Biochem.* **62,** 401–406 (2015).
27. Yamada, D. *et al.* Role of crosstalk between interleukin-6 and transforming growth factor-beta 1 in epithelial-mesenchymal transition and chemoresistance in biliary tract cancer. *Eur. J. Cancer* **49,** 1725–1740 (2013).
28. Yao, Z. *et al.* TGF-β IL-6 axis mediates selective and adaptive mechanisms of resistance to molecular targeted therapy in lung cancer. *Proc. Natl Acad. Sci. USA* **107,** 15535–15540 (2010).
29. Russo, J. E. & Hilton, J. Characterization of cytosolic aldehyde dehydrogenase from cyclophosphamide resistant L1210 cells. *Cancer Res.* **48,** 2963–2968 (1988).
30. Tam, W. L. & Weinberg, R. A. The epigenetics of epithelial–mesenchymal plasticity in cancer. *Nature Med.* **19,** 1438–1449 (2013).
31. Trimboli, A. J. *et al.* Direct evidence for epithelial–mesenchymal transitions in breast cancer. *Cancer Res.* **68,** 937–945 (2008).
32. Yu, M. *et al.* Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition. *Science* **339,** 580–584 (2013).
33. Korpal, M. *et al.* Direct targeting of Sec23a by miR-200s influences cancer cell secretome and promotes metastatic colonization. *Nature Med.* **17,** 1101–1108 (2011).
34. Gupta, P. B. *et al.* Identification of selective inhibitors of cancer stem cells by high-throughput screening. *Cell* **138,** 645–659 (2009).
35. Diessner, J. *et al.* Targeting of preexisting and induced breast cancer stem cells with trastuzumab and trastuzumab emtansine (T-DM1). *Cell Death Dis.* **5,** e1149 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported by a grant from the US Department of Defense CDMRP LCRP (LC110643). K.F. is supported by a fellowship from the NIH (1 F31 CA186510-01). V.M. was supported by the National Cancer Institute sub-award (U54 CA149196-05) and WCMC Meyer Cancer Center Pilot Funding. This work was also supported by funds from The Neuberger Berman Foundation Lung Cancer Research Center; the Arthur and Myra Mahon Donor-Advised Fund; the Liz Claiborne and Art Ortenberg Foundation; the Douglas & Katherine McCormick Family Foundation, the R. & M. Goldberg Family Foundation; the P. & C. Collins Fund; the Eliot Stewart 'Wren' Fund; the William and Shelby Modell Family Foundation Trust; and generous funds donated by patients in the Division of Thoracic Surgery to N.K.A. The funding organizations played no role in experimental design, data analysis or manuscript preparation.

Author Contributions D.G., K.R.F. and V.M. designed the experiments. K.R.F. and D.G. performed the experiments. A.D., S.L., H.C., T.E.R. and S.R. provided technical support with experiments and animal work. L.T.V. and N.K.A. made critical comments to improve the study design. J.S., F.L. and S.T.C.W. performed RNA-sequencing analysis. J.T. and R.F.S. generated the Vim-Cre transgenic mice. K.R.F., D.G. and V.M. wrote and edited the manuscript with input from the other authors. All authors discussed the results and conclusions drawn from them.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to V.M. (vim2010@med.cornell.edu) or D.G. (dig2009@med.cornell.edu).

METHODS

Animals. Wild-type C57BL/6 and FVB/n mice, and transgenic mice with ACTB–tdTomato–eGFP (stock no. 007676), *Fsp1*–Cre (stock no. 012641), MMTV–PyMT (stock no. 002374), and MMTV–Neu (stock no. 002376) were obtained from The Jackson Laboratory. The vimentin–CreER mouse was a kind gift from the laboratory of R. F. Schwabe at Columbia University. CB-17 SCID mice were obtained from Charles River Laboratories. All mouse strains obtained were bred in the animal facility at Weill Cornell Medical College. All animal work was conducted in accordance with a protocol approved by the Institutional Animal Care and Use Committee at Weill Cornell Medical College.

The ACTB–tdTomato–EGFP and *Fsp1*–Cre mice were bred together to obtain double transgenic mice and then bred with MMTV–PyMT or MMTV–Neu mice to obtain the tri-PyMT and tri-Neu triple-transgenic mice, respectively. Double transgenic male mice carrying ACTB–tdTomato–eGFP and MMTV–PyMT were crossed with the vimentin–CreER mice to obtain the tri-PyMT/Vim triple-transgenic mice. Genotyping for each transgenic line was performed following the standardized protocols as described in the website of The Jackson Laboratory. Genotyping for vimentin–CreER was done using forward primer 5′-CCCCTTCCTCACTTCTTTCC and reverse primer 5′-ATGTTTAGCTGGCCCAAATG.

Tamoxifen injection. To induce vimentin–CreER activity in the tri-PyMT/Vim mice, Tamoxifen (Sigma-Aldrich, 2 mg per mouse, dissolved in corn oil) was administered through intraperitoneal injections, three times per week starting when the primary tumours appear (at 8 weeks of age) and continuing for 6 weeks until metastasis developed in the lung.

Establishing tri-PyMT cell line. The primary tumour of the tri-PyMT mouse (12-week-old female) was surgically removed under sterile conditions. Tumour tissue was sliced into ~1 mm^3 blocks and implanted into the fat pad (no. 4 on the right side) of CB-17 SCID mice. The secondary tumour was used to establish the tri-PyMT cell line, eliminating the contamination of fluorescent positive stromal cells in the tumour tissue from tri-PyMT transgenic mice.

Tumour tissue was minced and digested with an enzyme cocktail (Collagenase A, elastase, and DNase I, Roche Applied Science) in HBSS buffer at 37 °C for 30 min. The cell suspension was strained through a 40-μm cell strainer (BD Biosciences). Cells were washed with PBS three times and uploaded in the Aria III cell sorter (BD Biosciences). The sorted RFP^+ cells were cultured in DMEM supplemented with 10% fetal bovine serum. The PyMT oncogene expression in the established cell line was confirmed by RT–PCR (Extended Data Fig. 4c). The tumorigenic ability of these cells was confirmed throughout the study.

To determine EMT induction by TGF-β, cells were cultured for one week in DMEM with 2% FBS and 2 ng ml^{-1} TGF-β1 (R&D Systems). The GFP^+ cell ratio was quantified by flow cytometry.

To generate the miR-200 overexpressing cell line, a pLenti 4.1 Ex miR-200b-200a-429 construct[20], was obtained from Addgene. To eliminate the contamination of fluorescent marker expression in targeted cells, the GFP gene in this construct was removed by BstBI/XbaI digestion followed by blunted self-ligation. Lentivirus was packaged by co-transfection of the pLenti-miR-200 construct and packaging plasmids into HEK293T cells. tri-PyMT cells (passage 2) were infected with the lentivirus. Infected cells (tri-PyMT miR-200) were selected by culturing with puromycin (2 μg ml^{-1}) for 14 days. A control tri-PyMT cell line was generated by infecting cells with lentivirus carrying the puromycin resistance gene, following the same procedure in parallel.

Orthotopic breast tumour model. To establish an orthotopic breast tumour model, we first purified RFP^+ cells from passages 10–15 of tri-PyMT cell culture by FACS. The purified RFP^+ tri-PyMT cells (1×10^6 cells with purity >99%, Extended Data Fig. 5a) were injected into the mammary fat pad of 8-week-old female CB-17 SCID mice. The growth of the primary tumour was monitored by external calliper measurement once a week. In approximately 4 weeks, the primary tumour was surgically removed and the incision was closed with wound clips. The tumour size did not exceed 5% of total body weight as permitted in the IACUC protocol. Animals were euthanized 4 weeks after primary tumour removal to analyse the development of pulmonary metastasis. For animals subjected to chemotherapy, Cyclophosphamide (CTX, Sigma-Aldrich, 100 mg kg^{-1}) was administered once per week, for 2 weeks prior and 2 weeks after surgery.

Tissue processing, immunofluorescence and microscopy. The harvested primary tumours and PBS-perfused lungs bearing metastases were fixed in 4% paraformaldehyde overnight, followed by 30% sucrose for 2 days, and then embedded in Tissue-tek O.C.T. embedding compound (Electron Microscopy Sciences). Serial sections (10 μm, at least 10 sections) were prepared for histological analysis by haematoxylin and eosin staining, and immunofluorescent staining following standardized protocols.

Primary antibodies used in this study include CD45 (30-F11, BioLegend), E-cadherin (DECMA-1, BioLegend), vimentin (sc-7557, Santa Cruz), PyMT (ab15085, Abcam), Neu (sc-284, Santa Cruz), Ki67 (ab15580, Abcam), and active caspase-3 (C92-605, BD Pharmingen). Primary antibodies were directly conjugated to Alexa Fluor 647 using an antibody labelling kit (Invitrogen) performed as per manufacturer's instructions and purified over BioSpin P30 columns (Bio-Rad). GFP^+ and RFP^+ cells were detected by inherent fluorescence.

Fluorescent images were obtained using a computerized Zeiss fluorescent microscope (Axiovert 200M), fitted with an apotome and an HRM camera. Images were analysed using Axiovision 4.6 software (Carl Zeiss).

Flow cytometry and cell sorting. For the metastatic lungs and primary tumours, cell suspensions were prepared by digesting tissues with an enzyme cocktail (collagenase A, elastase, and DNase I, Roche Applied Science) in HBSS buffer at 37 °C for 30 min. For cultured cells, cells were collected through trypsinization. A single-cell suspension was prepared by filtering through a 30-μm cell strainer (BD Biosciences). Then cells were stained following a standard immunostaining protocol. In brief, cells were pre-blocked with 2% FBS plus Fc block (CD16/CD32, 1:30, BD Biosciences) and then incubated with the primary antibody against E-cadherin (DECMA-1, BioLegend). SYTOX Blue (Invitrogen) was added to the staining tube in the last 5 min to facilitate the elimination of dead cells. GFP^+ and RFP^+ cells were detected by their intrinsic signals. The stained samples were analysed using the LSRII flow cytometer coupled with FACS Diva software (BD Biosciences). Flow cytometry analysis was performed using a variety of controls including isotype antibodies, unstained and single-colour stained samples for determining appropriate gates, voltages and compensations required in multivariate flow cytometry.

For sorting live cells back for further culturing or injection into animals, we used the Aria II cell sorter coupled with FACS Diva software (BD Biosciences). The preparation of cells for sorting was performed under sterile conditions. The purity of subpopulations after sorting was confirmed by analysing post-sort samples in the sorter again.

Quantitative RT–PCR analysis. Total RNA was extracted by using the RNeasy Kit (Qiagen), and miRNA via the mirVana miRNA isolation kit (Life Technologies), and converted to cDNA using qScript cDNA SuperMix (Quanta Biosciences) and RT–PCR. qPCR was performed with the appropriate primers (sequences shown in the table) and iQ™ SYBR Green master mix (Bio-Rad). PCR protocol: initial denaturing at 95 °C for 3 min, 40 cycles of 95 °C for 20 s, 60 °C for 30 s, and 72 °C for 30 s, followed by final extension at 72 °C for 5 min and melt curve analysis was applied on a Bio-Rad CFX96 Real Time System (Bio-Rad) coupled with Bio-Rad-CFX Manager software. Primers used are as follows: *GAPDH*, forward, 5′-GGTCCTCAGTGTAGCCCAAG-3′; reverse 5′-AATGTGTCCGTCGTGGATCT-3′; *Cdh1* (E-cadherin), forward, 5′-ACACCGATGGTGAGGGTACACAGG-3′; reverse, 5′-GCCGCCACACACAGCATAGTCTC-3′; *Ocln*, forward, 5′-TGCTAAGGCAGTTTTGGCTAAGTCT-3′, reverse, 5′-AAAAACAGTGGTGGGGAACGTG-3′; *Vim*, forward, 5′-TGACCTCTCTGAGGCTGCCAACC-3′; reverse, 5′-TTCCATCTCACGCATCTGGCGCTC-3′; *Cdh2* (N-cadherin), forward, 5′-AAAGAGCGCCAAGCCAAGCAGC-3′; reverse, 5′-TGCGGATCGGACTGGGTACTGTG-3′; *FSP-1*, forward, 5′-CCTGTCCTGCATTGCCATGAT-3′, reverse, 5′-CCCACTGGCAAACTACACCC-3′; *Snai1*, forward, 5′-ACTGGTGAGAAGCCATTCTCCT-3′; reverse, 5′-CTGGCACTGGTATCTCTTCACA-3′; *Snai2*, forward, 5′-TTGCAGACAGATCAAACCTGAG-3′; reverse, 5′-TGTTTATGCAGAAGCGACATTC-3′; *Twist1*, forward, 5′-AGCTACGCCTTCTCCGTCTG-3′; reverse, 5′-CTCCTTCTCTGGAAACAATGACA-3′; *Zeb-1*, forward, 5′-GATTCCCCAAGTGGCATATACA-3′; reverse, 5′-TGGAGACTCCTTCTGAGCTAGTG-3′; *Zeb-2*, forward, 5′-TGGATCAGATGAGCTTCCTACC-3′; reverse, 5′-AGCAAGTCTCCCTGAAATCCTT-3′; *PyMT*, forward, 5′-ACTGCTACTGCACCCAGACA-3′; reverse, 5′-CTGGAAGCCGGTTCCTCCTA-3′; *GFP*, forward, 5′-CCACATGAAGCAGCACGACT-3′; reverse, 5′-GGGTCTTGTAGTTGCCGTCG-3′; *RFP*, forward, 5′-AGCGCGTGATGAACTTCGAG-3′; reverse, 5′-CCGCGCATCTTCACCTTGTA-3′.

RNA-sequencing analysis. Total RNA was extracted from sorted RFP^+ and GFP^+ tri-PyMT cells with the RNeasy Kit (Qiagen). RNA-seq libraries was constructed and sequenced following standard protocols (Illumina). Single-end RNA-seq reads were mapped to UCSC mouse genome (GRCm38/mm10) using Tophat2. FPKM values for each gene were estimated by Cufflinks and statistical analysis was done using Cuffdiff2. Heat maps for differentially expressed genes with adjusted P values <0.05 were drawn using gplots R package.

Western blot analysis. Cells were homogenized in 1× RIPA lysis buffer (Millipore) with protease inhibitors (Roche Applied Science). Samples were boiled in 1× Laemmli buffer and 10% β-mercaptoethanol, and loaded onto 12% gradient Tris-glycine gels (Bio-Rad). Western blotting was performed using antibodies specific for E-cadherin (clone DECMA-1), vimentin (clone RV202, BD Pharmingen), and β-actin (clone AC-15, Sigma-Aldrich).

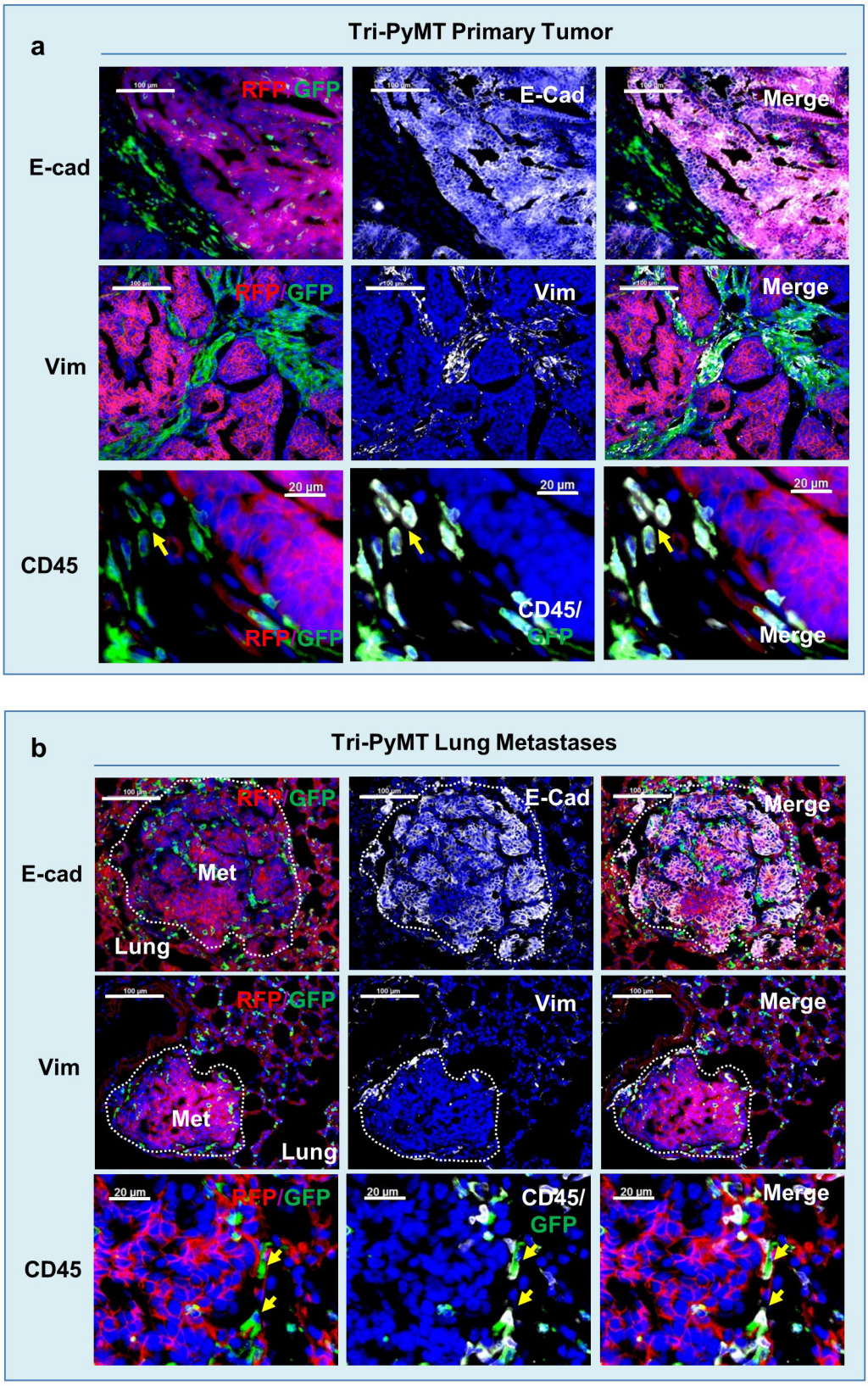
Cell apoptosis and viability assays. To determine apoptosis of RFP^+ and GFP^+ cells, tri-PyMT cells (Passage 10) were seeded on adherent six-well plates

(1×10^6 cells), and treated with 4-hydroperoxy cyclophosphamide (Santa Cruz) for 48 h. After treatment, cells were trypsinized and stained with APC-conjugated Annexin V (BD Biosciences) and SYTOX Blue (Invitrogen) for apoptotic-cell labelling. The stained cells were analysed in the LSRII flow cytometer to quantify the percentage of apoptotic, dead, and live RFP^+ and GFP^+ cells by FACS Diva software. To determine the viability of tri-PyMT control and miR-200-expressing cells treated with CTX, cells were plated in 96-well adherent black-walled plates (1×10^4 cells), and treated with 4-hydroperoxy cyclophosphamide for 48 h. After treatment, cell viability was measured with the CellTiter-Glo Luminescent Cell Viability Assay (Promega).

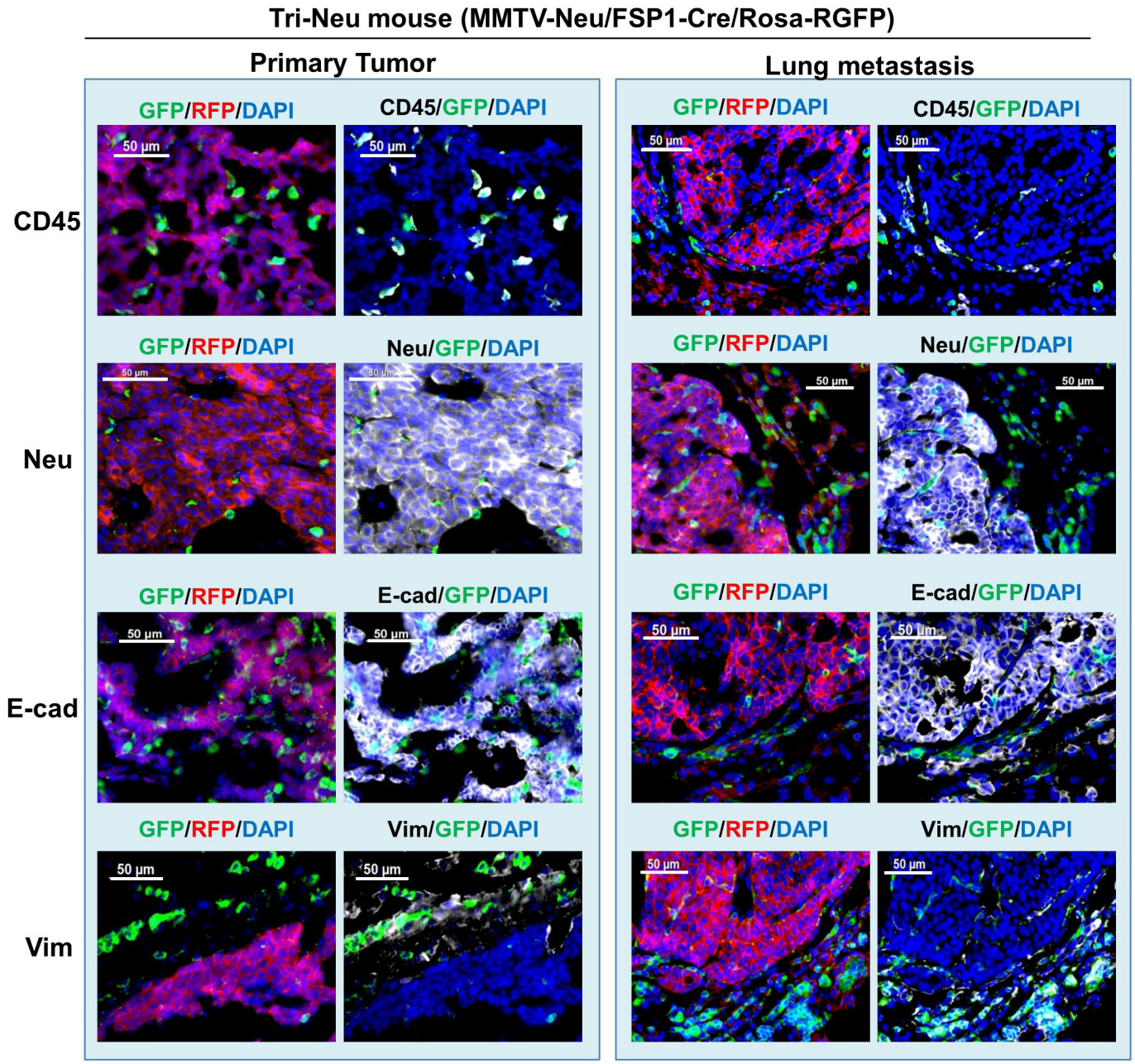
Cell migration assay. 1×10^5 tri-PyMT cells were seeded in a six-well plate. Real-time images of cells (including phase, GFP and RFP channels) were taken under a computerized Zeiss microscope (Axiovert observation) every 10 min for 10 h. Movement of individual cells (>10 RFP^+ and >10 GFP^+ cells in each field, >2 fields were analysed) were tracked with ImageJ software, and the distance that was travelled during that time was measured as indicated.

ALDH activity assay. RFP^+ and GFP^+ tri-PyMT cells (1×10^6 cells each) were freshly sorted from culture by FACS and then homogenized in cold ALDH Assay buffer provided in the ALDH Activity Colorimetric Assay Kit (Biovision Inc.) Following the protocol, ALDH substrate and acetaldehyde were added. ALDH activities in samples were measured by OD at 450 nm in kinetic mode (every 3 min for 60 min).

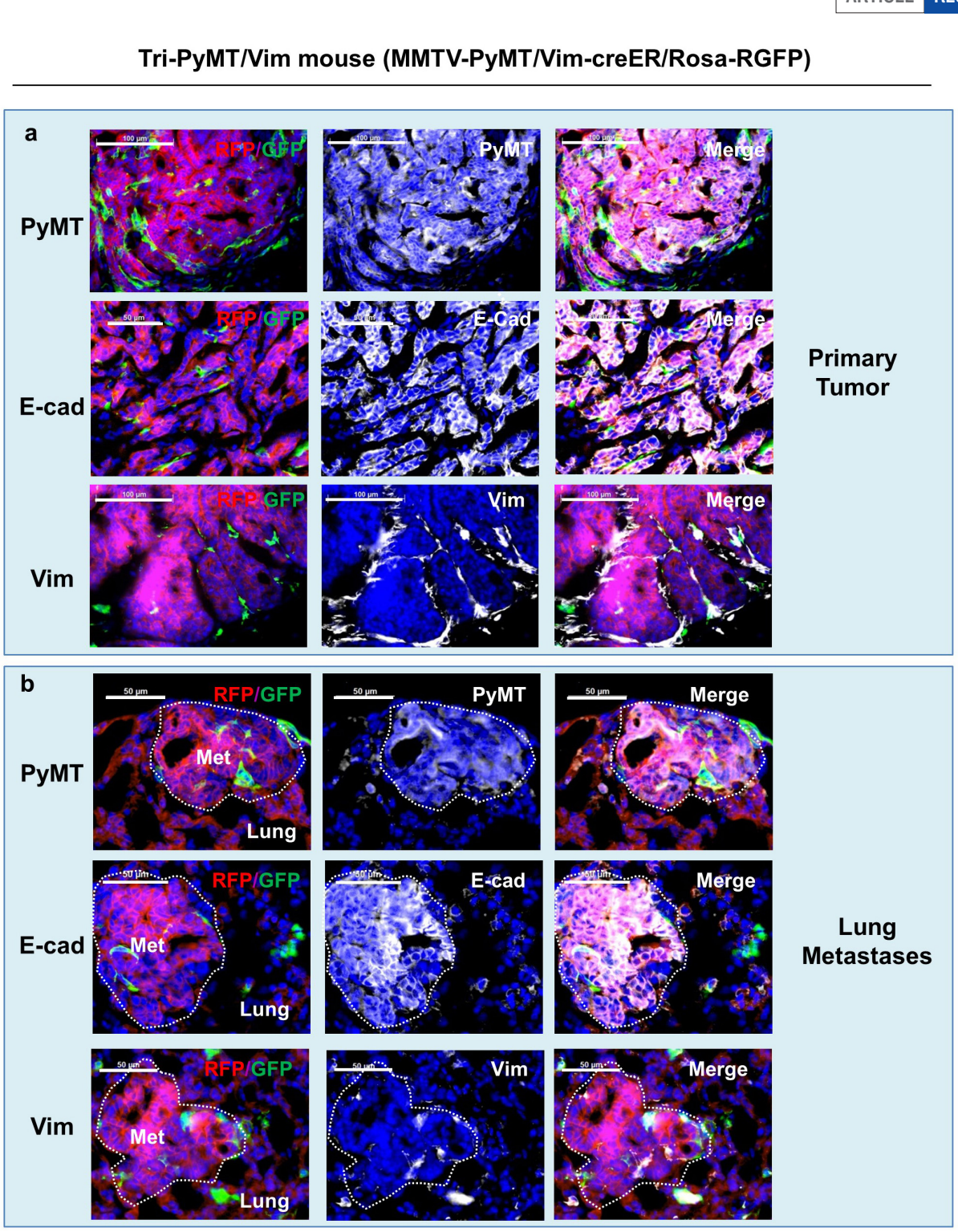
Statistical analysis. To determine the sample size of animal experiments, we used power analysis assuming $\frac{\text{difference in means}}{\text{standard deviation}} > 2.5$. Therefore, all animal experiments were conducted with ≥ 5 mice per group to ensure adequate power between groups by two-sample *t*-test comparison. Animals were randomized within each experimental group. No blinding was applied in performing experiments. Results are expressed as mean $\pm$ s.e.m. Data distribution in groups and significance between different treatment groups was analysed by using the Mann–Whitney *U*-test in GraphPad Prism software. *P* values <0.05 were considered significant. Error bars depict s.e.m., except where indicated otherwise.



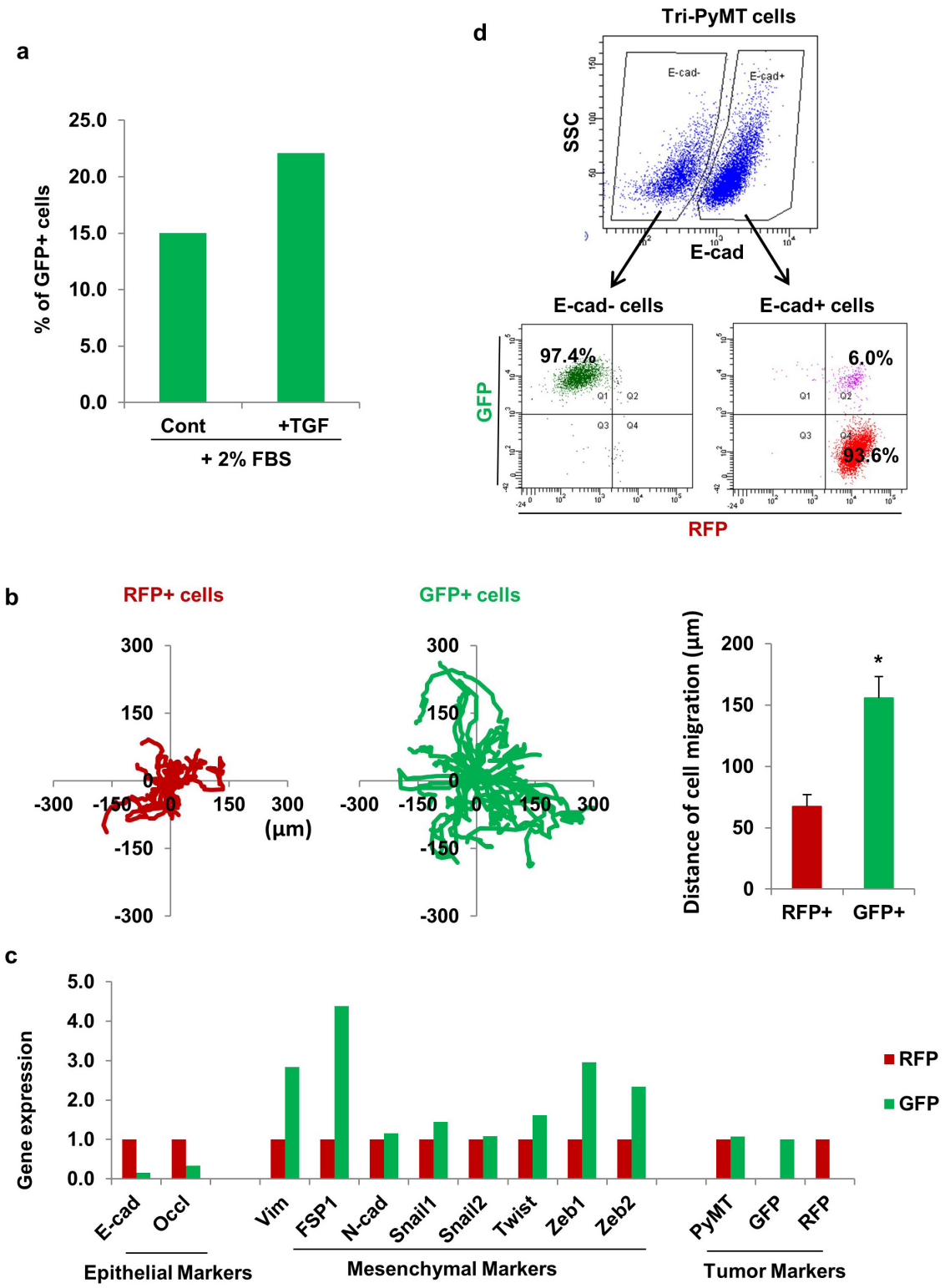
Extended Data Figure 1 | Characterization of the primary tumour and lung metastasis of tri-PyMT mice. a, **b**, Sections of primary tumours (**a**) and lungs (**b**) from tri-PyMT mice were immunostained for E-cadherin (E-cad, top), vimentin (Vim, middle) and CD45 (bottom) in white pseudo-colour. Representative images are shown ($n > 5$ mice). Note the co-localization of PyMT with RFP, and CD45 with GFP (as indicated by arrows), in both primary tumours and lung metastases.



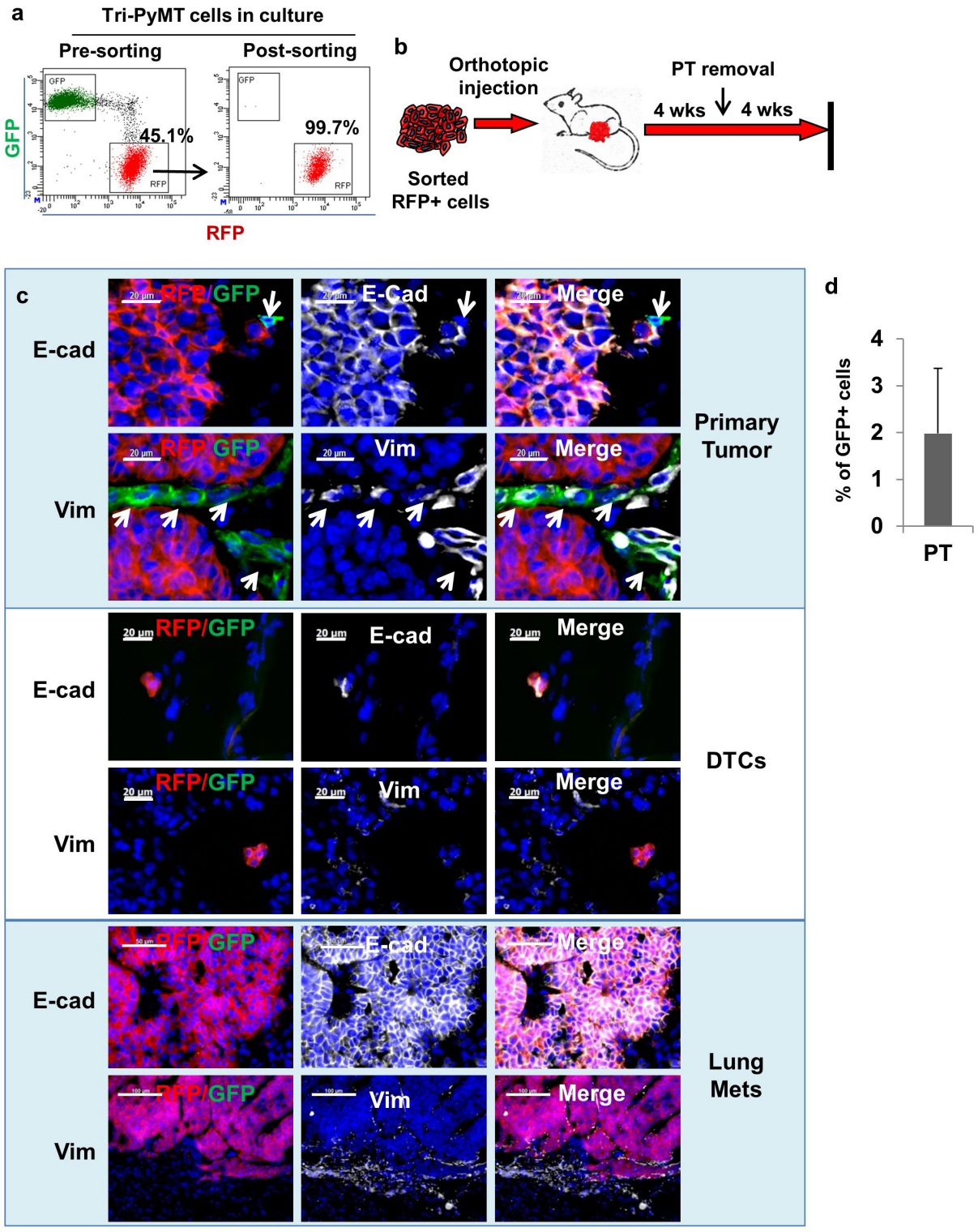
Extended Data Figure 2 | Characterization of the primary tumour and lung metastasis of tri-Neu mice. Sections of primary tumours (left panel) and lungs (right panel) from tri-Neu mice were immunostained for CD45, Neu, E-cadherin, and vimentin (in white pseudo-colour). Representative images are shown ($n > 5$ mice). Note that both primary tumours and lung metastases are largely composed of epithelial RFP^+ tumour cells.



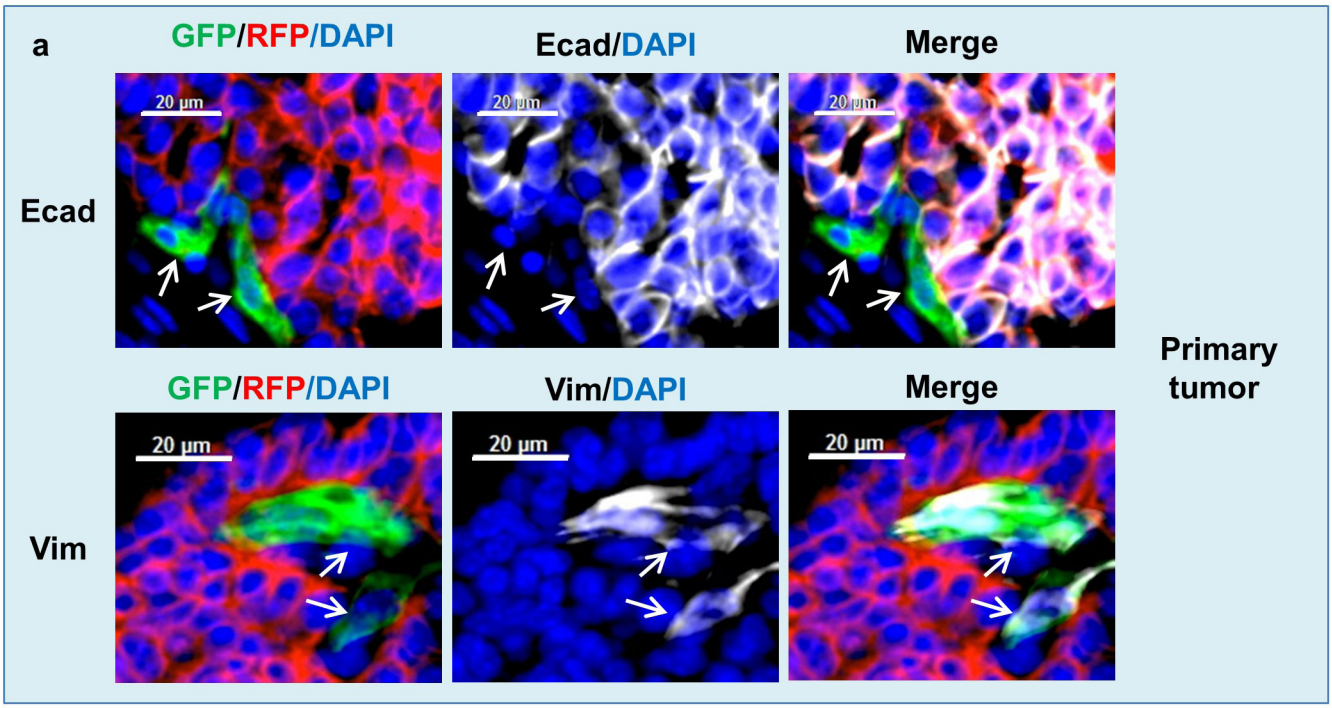
Extended Data Figure 3 | Characterization of the primary tumour and lung metastasis of tri-PyMT/Vim mice. Tri-PyMT/Vim mice were obtained by crossing MMTV–PyMT, vimentin–Cre and Rosa26–RFP–GFP transgenic mice. **a**, **b,** Sections of primary tumours (**a**) and lungs (**b**) were immunostained for PyMT, E-cadherin and vimentin (in white pseudo-colour). Representative images are shown ($n > 5$ mice). Note that both primary tumours and lung metastases are largely composed of epithelial RFP^+ tumour cells.

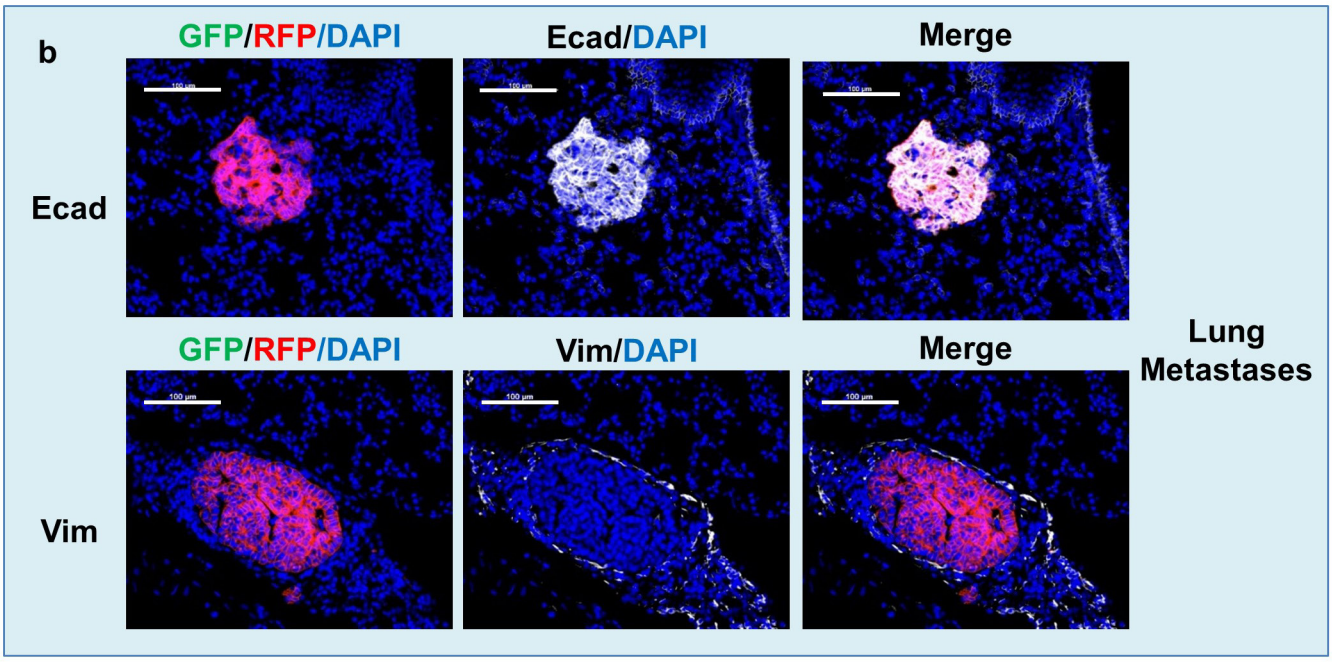


Extended Data Figure 4 | Characterization of tri-PyMT cells. a, EMT of tri-PyMT cells with TGF. RFP^+ tri-PyMT cells were sorted by flow cytometry and cultured in medium containing 2% FBS with or without TGF-β1 (2 ng ml^{-1}) for 3 days. Plot shows quantification of the percentage of GFP^+ cells analysed by flow cytometry ($n = 2$ biological replicates). **b**, Cell migration assay of tri-PyMT cells. The tracing plots show the movement of individual RFP^+ and GFP^+ cells in 10 h of live imaging. Quantification plot (right panel) showed the average distance that RFP^+ and GFP^+ cells have moved during the time frame ($n > 20$, $^*P < 0.01$). **c**, Relative expression of epithelial, mesenchymal and tumour markers in sorted RFP^+ and GFP^+ tri-PyMT cells as determined by qRT–PCR with *Gapdh* as the internal control. $n = 2$ individual experiments. **d**, EMT of tri-PyMT cells is reported by fluorescent marker switch. Flow cytometry plot shows E-cadherin$^-$ (E-cad$^-$) and E-cadherin$^+$ (E-cad$^+$) subpopulations of tri-PyMT cells (upper panel). Of the E-cad$^-$ and E-cad$^+$ subsets, the populations were further dissected according to innate fluorescence (lower panel). Numbers indicate the percentage of GFP^+, RFP^+, or transitioning (Q2) cells in the parental E-cad$^-$ or E-cad$^+$ subsets, respectively.



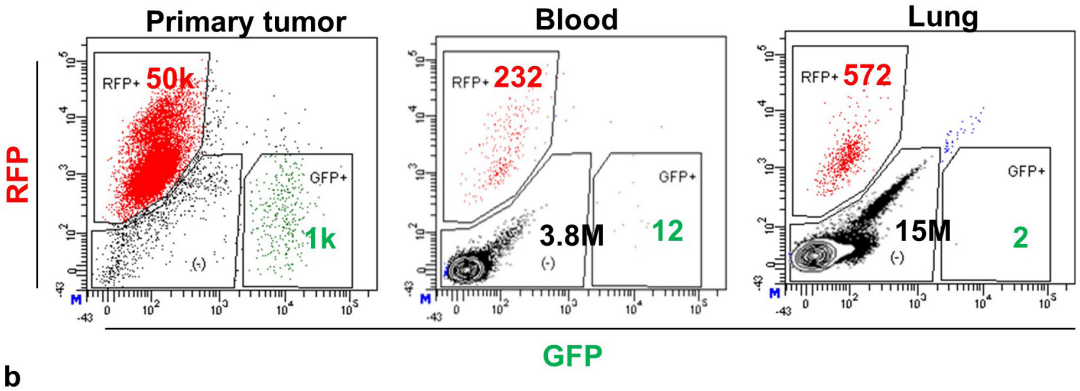
Extended Data Figure 5 | Establishing an orthotopic model with sorted RFP$^+$ tri-PyMT cells. a, Flow cytometry plots show tri-PyMT cells before and after sorting for RFP$^+$ cells. Numbers indicate the percentage and purity of RFP$^+$ cells used for establishing orthotopic breast tumours in mice. **b**, Schematic of the orthotopic breast tumour model with sorted RFP$^+$ tri-PyMT cells. Cells are injected into the mammary gland of wild-type mice to generate primary breast tumours, resection of primary tumour at 4 weeks and lung metastases evaluation in another 4 weeks. **c**, Characterization of tumour cells in the primary tumour, disseminated tumour cells (DTCs) and tumour cells in the lung metastasis of the tri-PyMT orthotopic model. Sections of primary tumours and lungs from tri-PyMT orthotopic mice were immunostained for E-cadherin and vimentin (in white pseudo-colour). Essentially all RFP$^+$ tumour cells are detected as E-cad$^+$/Vim$^-$, while the scattered GFP$^+$ tumour cells in the primary tumour are E-cad$^-$/Vim$^+$ (as indicated by arrows in the top panel). Representative images are shown ($n = 8$). **d**, Plot shows the percentage of GFP$^+$ cells out of total tumour cells (GFP$^+$ plus RFP$^+$, $n = 6$).





Extended Data Figure 6 | Characterization of EMT status of orthotopic tri-Vim–PyMT primary tumours. a, **b**, Sections of tri-Vim–PyMT orthotopic primary tumours (**a**) and metastatic lung (**b**) were immunostained for E-cadherin and vimentin (in white pseudo-colour). As expected, RFP^+ tumour cells are entirely E-cadherin-positive and vimentin-negative, GFP^+ tumour cells are vimentin-positive and E-cadherin-negative, and lung metastases are epithelial and RFP^+.

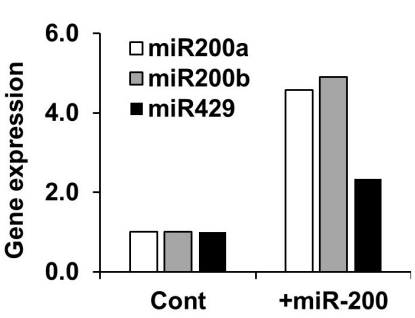
a



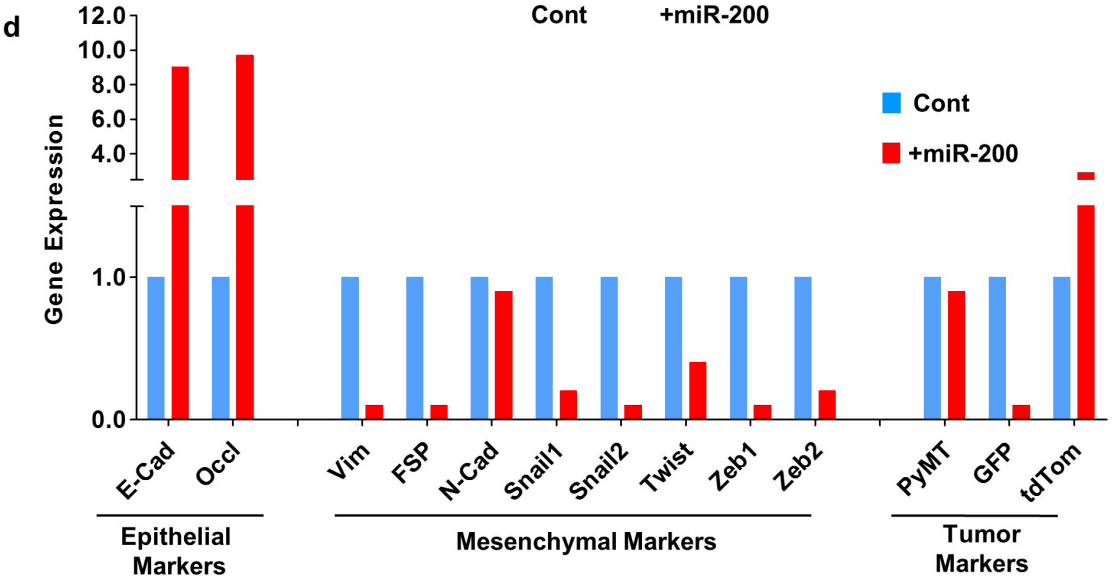
b

RFP : GFP	Primary tumor	Blood	Lung
Mouse 1	44 : 1	19 : 1	286 : 1
Mouse 2	181 : 1	12 : 1	296 : 0
Mouse 3	84 : 1	11 : 1	45 : 1
Mouse 4	120 :1	15 : 1	110 : 1

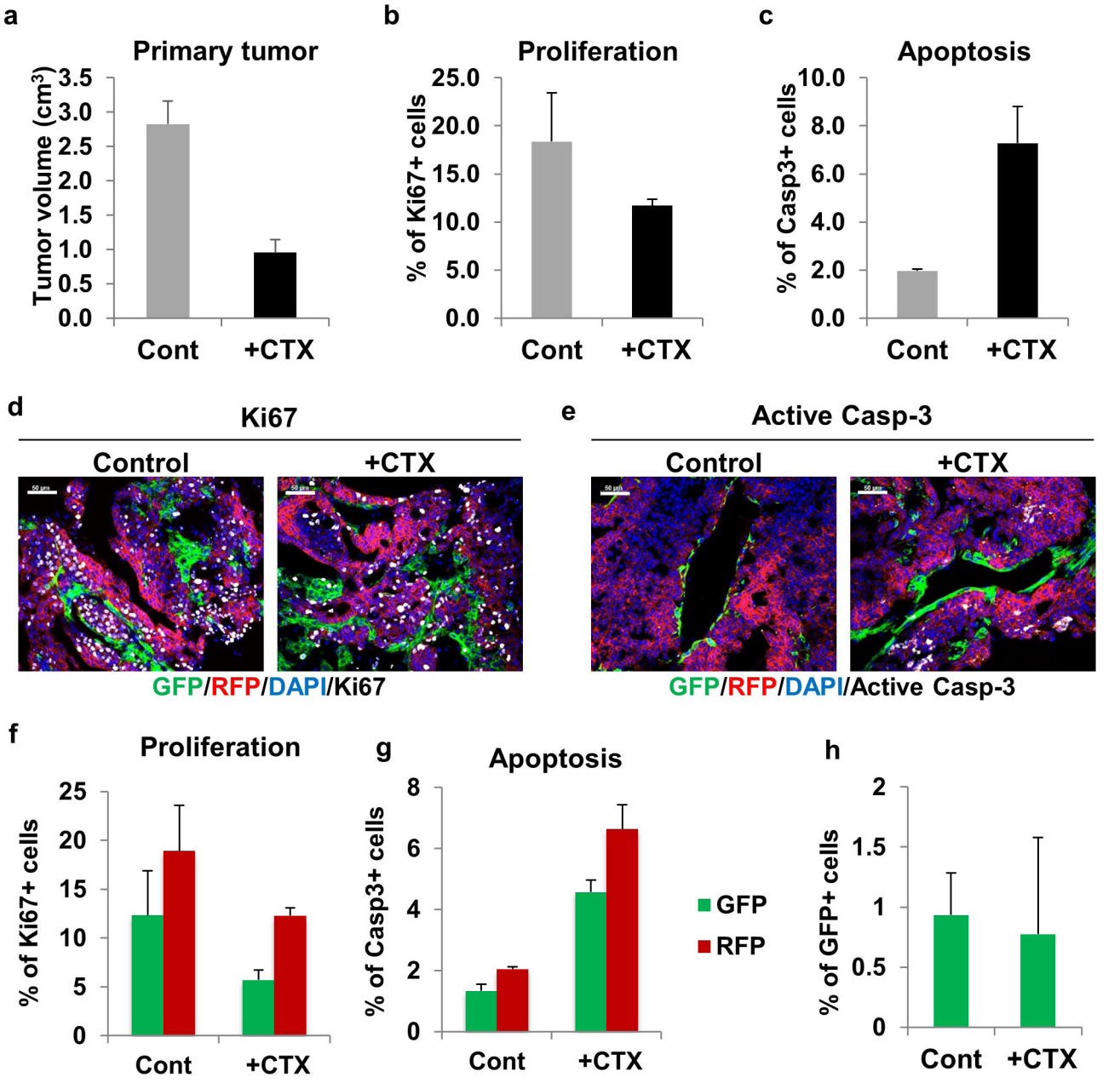
c



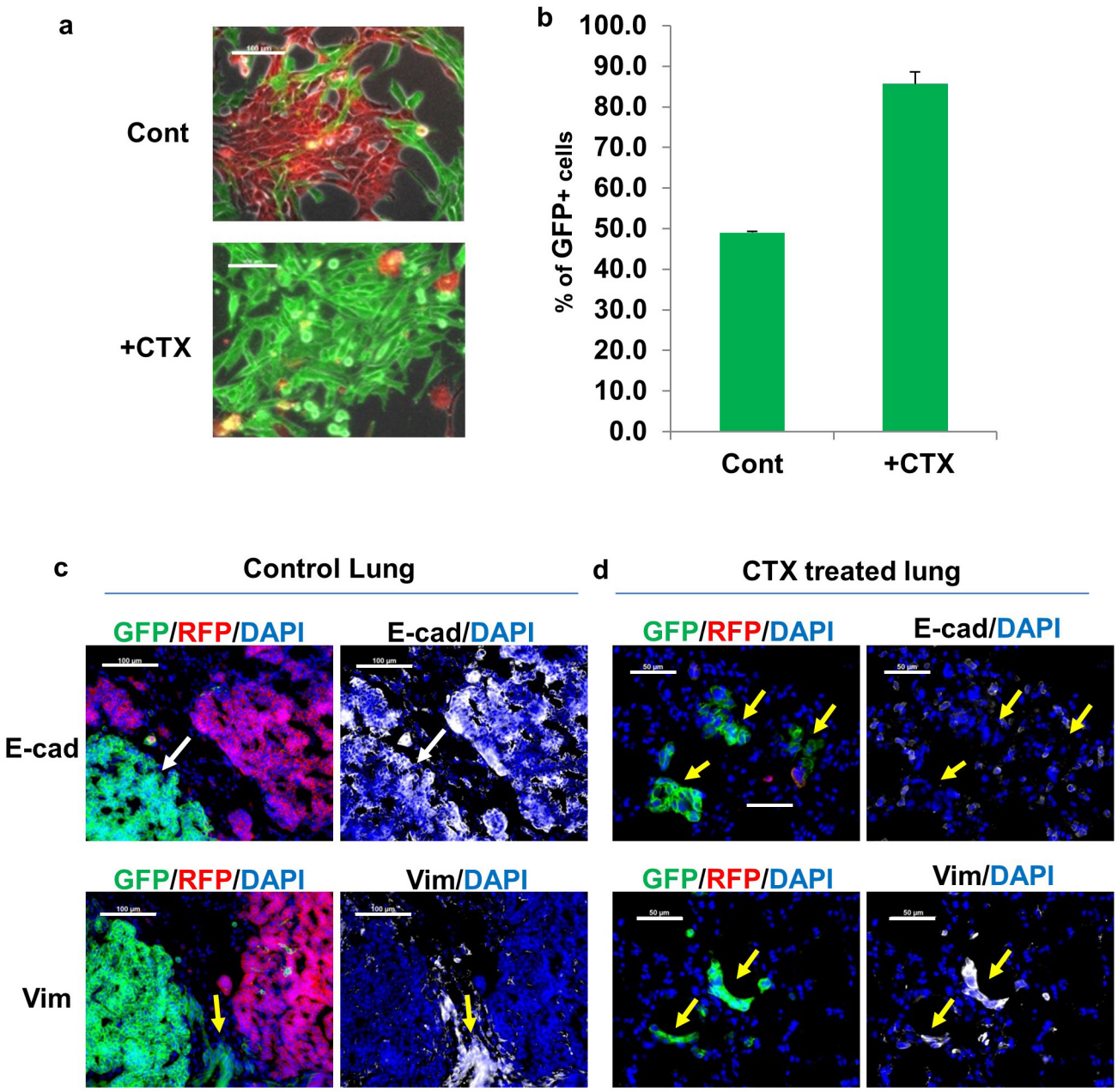
d



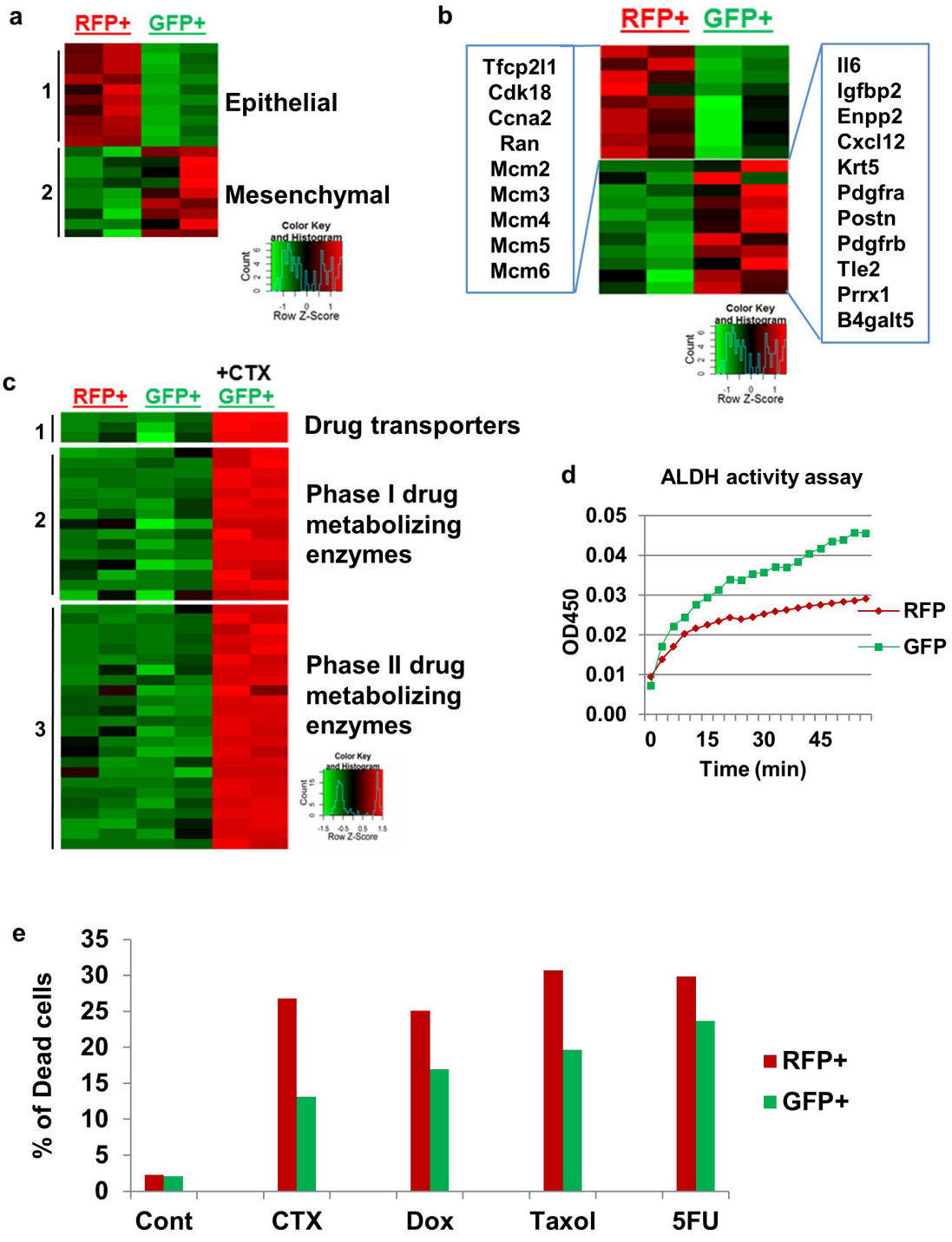
Extended Data Figure 7 | Dissemination of tri-PyMT cells *in vivo*. **a**, Disseminated tumour cells are RFP$^+$ and epithelial. RFP$^+$ tri-PyMT cells were injected into the fat pad of mice. The fluorescence of the primary tumour, circulating tumour cells in the blood and disseminated tumour cells in the lung were analysed by flow cytometry. The flow cytometry plots depicted are the enumeration of RFP$^+$ and GFP$^+$ cells. **b**, The ratios of detected RFP$^+$ versus GFP$^+$ cells are shown in the chart ($n = 4$ mice). **c**, Relative expression of miR-200-family microRNAs in tri-PyMT control and miR-200-expressing cells. $n = 2$ individual experiments. **d**, Relative expression of EMT markers and tumour markers in tri-PyMT control and mir-200-expressing cells as determined by qRT–PCR with *Gapdh* as the internal control, $n = 2$ individual experiments.



Extended Data Figure 8 | Effects of CTX therapy on primary tumours. **a**, Quantification of primary tumour growth after 2 weeks of CTX therapy. For tumour growth data see accompanying Source Data. **b**, Proliferation status of primary tumour cells as detected by Ki67 staining in control mice and after 2 weeks of CTX therapy. **c**, Level of apoptosis in primary tumours as detected by active caspase-3 staining in control mice and after 2 weeks of CTX therapy. **d**, **e**, Representative images of Ki67 (**d**) and active caspase-3 staining (**e**) (white pseudo-colour) of primary tumours in control mice and CTX-treated mice. Scale bars, 50 μm. **f**, Proliferation status of RFP^+ and GFP^+ primary tumour cells as detected by Ki67 staining in control and CTX-treated mice. **g**, Level of apoptosis in RFP^+ and GFP^+ primary tumours as detected by active caspase-3 staining in control and CTX-treated mice. **h**, Percentage of GFP^+ tumour cells in control and CTX-treated primary tumours. $n = 3$ mice for all figures described above. Quantification performed using ImageJ software.



Extended Data Figure 9 | EMT tumour cells are resistant to CTX treatment both *in vitro* and *in vivo*. **a**, **b**, Long-term CTX treatment *in vitro* results in a GFP^+ population. Tri-PyMT cells were subjected to 2 weeks cyclophosphamide (+CTX) treatment (4 μM). Fluorescent imaging (**a**) and flow cytometry (quantified, **b**, $n = 3$) exhibit the percentage of GFP^+ cells in the CTX-treated culture compared to untreated control cells **c**, **d**, EMT status of lung nodules in competitive survival assay. Representative fluorescent images of tri-PyMT lung metastases in untreated control lungs (**c**) and CTX-treated lungs (**d**), depicting RFP^+ and GFP^+ tumour cells. Immunostaining showing E-cadherin (E-cad) or vimentin (Vim) in white pseudo-colour. White arrow indicates GFP^+ tumour cells with epithelial phenotypes ($E\text{-}cad^+/Vim^-$), while the yellow arrow indicates GFP^+ cells with mesenchymal phenotypes ($E\text{-}cad^-/Vim^+$). Nuclei were counter-stained with DAPI. $n = 5$.



Extended Data Figure 10 | Gene expression profile analysis of RFP$^+$ and GFP$^+$ tri-PyMT cells. RFP$^+$ and GFP$^+$ tri-PyMT cells were sorted by flow cytometry and subjected to transcriptomic analysis by RNA-sequencing. **a**, Heat map of differentially expressed genes (adjusted $P < 0.05$) from RNA-seq of sorted RFP$^+$ and GFP$^+$ tri-PyMT cells, biologically duplicated. Genes that are established epithelial markers (Group 1) include *Cdh1* (which encodes E-cad), *Dsp*, *Epcam*, *Fgfbp1*, *Krt18*, *Krt19*, *Ocln*, *Tjp3*, *Krt14* and *Tjp2*; the mesenchymal markers (Group 2) include *Cdh2* (which encodes N-cad), *Col23a1*, *Col3a1*, *Col5a1*, *Col6a2*, *Fsp1*, *Mmp3*, *Wnt5a* and *Zeb1*. **b**, Cell cycle (left panel) and chemoresistance-related (right panel) genes alternatively regulated in RFP$^+$ and GFP$^+$ cells. **c**, GFP$^+$ tri-PyMT cells were also sorted from CTX-treated (4 μM) samples. Interestingly, a branch of genes related to drug metabolism were significantly elevated in CTX-treated GFP$^+$ cells. Group 1 genes are drug transporters including *Abcb1a*, *Abcb1b* and *Abcc1*. Group 2 genes are phase I drug-metabolizing enzymes including *Adh7*, *Aldh1a1*, *Aldh1a3*, *Aldh1l1*, *Aldh1l2*, *Aldh2*, *Aldh3a1*, *Aldh3a2*, *Aldh3b2*, *Aldh4a1*, *Cyp1a1*, *Cyp2f2*, *Cyp2j6*, *Ptgs1* and *Ptgs2*. Group 3 genes are phase II drug metabolizing enzymes including *Aox1*, *Blvrb*, *Ces2e*, *Ces2f*, *Ces2g*, *Chst1*, *Ephx1*, *Fmo1*, *Gpx2*, *Gsta3*, *Gsta4*, *Gstm2*, *Gsto1*, *Gstp1*, *Gstt3*, *Maoa*, *Mgst1*, *Mgst2*, *Nat6*, *Nat9*, *Nqo1*, *Pon3*, *Ugt1a6a* and *Ugt1a7c*. **d**, Aldehyde dehydrogenase (ALDH) activity assay. Cell lysates were prepared from flow cytometry-sorted RFP$^+$ and GFP$^+$ tri-PyMT cells. ALDH activity in samples was measured by OD at 450 nm in a kinetic mode (every 3 min for 60 min). Representative result from two independent experiments depicted. **e**, EMT tumour cells (GFP$^+$ cells) showed resistance to multiple commonly used chemotherapies. Tri-PyMT cells were subjected to treatment with CTX (8 μM), doxorubicin (Dox, 2 μM), paclitaxel (Taxol, 10 μM) and fluorouracil (5FU; 1.6 μM) for 3 days. Flow cytometry analysis of apoptotic cells was performed after Annexin staining. The percentage of dead cells (Annexin$^+$) in RFP$^+$ and GFP$^+$ cells, respectively, was quantified. $n = 2$ biological replicates.

ARTICLE

doi:10.1038/nature15699

Allosteric ligands for the pharmacologically dark receptors GPR68 and GPR65

Xi-Ping Huang[1,2]*, Joel Karpiak[3]*, Wesley K. Kroeze[1]*, Hu Zhu[1]†, Xin Chen[4,5]†, Sheryl S. Moy[6], Kara A. Saddoris[6]†, Viktoriya D. Nikolova[6], Martilias S. Farrell[1]†, Sheng Wang[1], Thomas J. Mangano[1,2], Deepak A. Deshpande[7], Alice Jiang[1,2]†, Raymond B. Penn[7], Jian Jin[4,5]†, Beverly H. Koller[8], Terry Kenakin[1], Brian K. Shoichet[3] & Bryan L. Roth[1,2,5]

At least 120 non-olfactory G-protein-coupled receptors in the human genome are 'orphans' for which endogenous ligands are unknown, and many have no selective ligands, hindering the determination of their biological functions and clinical relevance. Among these is GPR68, a proton receptor that lacks small molecule modulators for probing its biology. Using yeast-based screens against GPR68, here we identify the benzodiazepine drug lorazepam as a non-selective GPR68 positive allosteric modulator. More than 3,000 GPR68 homology models were refined to recognize lorazepam in a putative allosteric site. Docking 3.1 million molecules predicted new GPR68 modulators, many of which were confirmed in functional assays. One potent GPR68 modulator, ogerin, suppressed recall in fear conditioning in wild-type but not in GPR68-knockout mice. The same approach led to the discovery of allosteric agonists and negative allosteric modulators for GPR65. Combining physical and structure-based screening may be broadly useful for ligand discovery for understudied and orphan GPCRs.

G-protein-coupled receptors (GPCRs)—the largest family of proteins encoded in the human genome—transduce signals for the most diverse endogenous ligands of any receptor family. Correspondingly, GPCRs are the most productive drug targets, with over 26% of US Food and Drug Administration (FDA)-approved drugs acting primarily through them. Astonishingly, of the 356 non-olfactory GPCRs, about 38% are understudied or 'orphan' receptors whose physiological roles, and often endogenous ligands, remain unknown[1]. Given the central role of GPCRs in physiology and disease, and the high conservation of orphan GPCRs among organisms from worms to humans, understudied and orphan GPCRs are probably functionally and therapeutically important. Indeed, for the few GPCRs deorphanized since 2003 (refs 1, 2 and http://www.guidetopharmacology.org/GRAC/FamilyDisplayForward?familyId=16), most have newly approved and investigational drugs[1,3]. As with kinases[4], epigenetic proteins[5] and proteases[6], ligands specific for orphan GPCRs will illuminate their biology and provide new areas for therapeutic intervention.

A key impediment to GPCR deorphanization is uncertainty about the proteins through which they signal, making functional assays problematic[1]. This difficulty is increased by the diverse ligands that GPCRs recognize, which range from protons and photons, small neurotransmitters and lipids, to peptides and folded proteins. Thus, generic functional screens are difficult for orphan GPCRs—one neither knows what class of compounds to screen, nor how to screen for it, much less how to demonstrate relevance—thereby explaining the slow progress in determining their roles in signalling and physiology[3].

GPR68 (also known as OGR1) exemplifies both the important roles these understudied and orphan receptors are thought to serve, and our difficulties in illuminating them. Together with GPR4, GPR65 and GPR132, GPR68 belongs to a family of proton-sensing GPCRs[7]. GPR68 couples to several signalling pathways through G_q, G_s, $G_{12/13}$ or $G_{i/o}$ proteins[7–10]. GPR68 is expressed in many tissues and has been implicated in many processes[11–16], but it is most abundant in mouse cerebellum[17] and hippocampus[11] (http://www.brain-map.org/), suggesting yet to be identified roles in brain function. In acidic microenvironments, GPR68 seems to regulate inflammatory processes in airway smooth muscle and other cells[18–20]. Surprisingly, studies with GPR68-knockout mice uncovered only modest changes in these functions[16,21,22]. Although GPR68 has been reported to be activated by a family of isoxazoles[15], their weak activity seems to be nonspecific[23,24] and could not be reproduced (see later). Thus, although GPR68 may have many roles, few of them are well-characterized by knockout and none is known in the central nervous system (CNS), where it is most highly expressed. Like other targets lacking small molecule reagents, GPR68 remains 'pharmacological dark matter'[1].

Here we describe an integrated experimental and computational approach to discover ligands that modulate GPR68. A lead compound that functions as a positive allosteric modulator (PAM) is demonstrated *in vitro* and *in vivo*, providing insights into GPR68 physiology. Application of the same approach found allosteric agonists and negative allosteric modulators for a second understudied GPCR, GPR65, suggesting that the approach may be broadly useful.

[1]Department of Pharmacology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, 27599-7365, USA. [2]National Institute of Mental Health Psychoactive Drug Screening Program (NIMH PDSP), School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7365, USA. [3]Department of Pharmaceutical Chemistry, University of California at San Francisco, Byers Hall, 1700 4th Street, San Francisco, California 94158-2550, USA. [4]Center for Integrative Chemical Biology and Drug Discovery (CICBDD), University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7363, USA. [5]Division of Chemical Biology and Medicinal Chemistry, Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7360, USA. [6]Department of Psychiatry and Carolina Institute for Developmental Disabilities (CIDD), University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7146, USA. [7]Center for Translational Medicine and Department of Medicine, Thomas Jefferson University, Philadelphia, Pennsylvania 19107, USA. [8]Department of Genetics, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7264, USA. †Present addresses: National Center for Advancing Translational Sciences (NCATs), 9800 Medical Center Drive, Rockville, Maryland 20850, USA (H.Z.); Department of Pharmacology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-7365, USA (X.C.); Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, Colorado 80309, USA (K.A.S.); Department of Genetics, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27514, USA (M.S.F.); Touro College of Osteopathic Medicine, 60 Prospect Avenue, Middletown, New York 10940, USA (A.J.); Department of Structural and Chemical Biology, Department of Oncological Sciences, Department of Pharmacology and Systems Therapeutics, Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA (J.J.). *These authors contributed equally to this work.

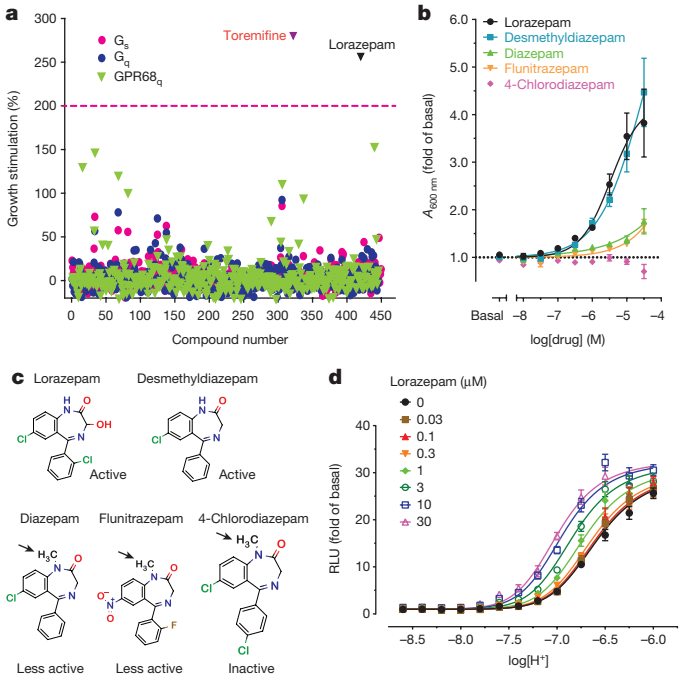


Figure 1 | Lorazepam is a GPR68-positive allosteric modulator. **a**, A library of approved drugs (10 μM) screened with yeast expressing chimaeric G_s, G_q or GPR68 and chimaeric G_q ($GPR68_q$) revealed lorazepam as a true and toremifine as a false positive. **b**, Concentration-dependent stimulation of GPR68 G_q-yeast growth by lorazepam and analogues. **c**, Structures of representative benzodiazepines (arrows denote methyl substituents that reduce GPR68 activity). **d**, Lorazepam is a GPR68-positive allosteric modulator for the agonist proton in the GPR68-mediated cAMP production. RLU, relative luminescence units. Data are mean ± s.e.m. of normalized results (**a**, **b**, **d**, $n = 3$) and concentration–response curves (**b**, **d**) were fit via a four-parameter logistic function (see Methods).

Yeast-based screen reveals GPR68 active compounds

In an initial campaign with 24 selected orphan and understudied GPCRs, we modified a yeast assay system[25] and screened a small library of approved drugs (http://www.nihclinicalcollection.com/ and Supplementary Fig. 1). We confirmed the known activity of short-chain carboxylic acids on the GPR41 and the GPR43 free fatty acid receptors (Extended Data Fig. 1a–d), and that of zinc (Extended Data Fig. 1e) and several other metals (Extended Data Fig. 1f–k) at GPR39. The most notable result was the finding that the benzodiazepine anxiolytic lorazepam was an agonist at GPR68 (Fig. 1).

Lorazepam activated GPR68 signalling, stimulating yeast growth by more than twofold (Fig. 1a). *N*-unsubstituted benzodiazepines were more efficacious than *N*-substituted benzodiazepines (Fig. 1b, c and Supplementary Table 1) and activated the receptor at both pH 6.5 and 7.4 (Extended Data Fig. 1l), with lorazepam most potently shifting the H^+ concentration–response profile (Fig. 1d and Extended Data Fig. 1m–p). The pH-dependence of lorazepam activity suggested that it functions as a PAM of GPR68; lorazepam did not affect the activity of the related receptors GPR4 or GPR65 (Extended Data Fig. 2a, b). When profiled against a panel of CNS targets, lorazepam had substantial activity only at the $GABA_A$ (γ-aminobutyric acid type A) receptor, its therapeutic target (Extended Data Fig. 3).

Modelling the GPR68–lorazepam complex

Little improvement in activity or selectivity was achieved by testing lorazepam analogues. This observation, and the potent $GABA_A$ receptor activity of the drug, led us to seek specific, optimizable molecules from computational docking screens of multi-million molecule libraries (Fig. 2).

We generated 407 homology 3D models for GPR68 templated on the CXCR4 structure (29% sequence identity, Extended Data Fig. 2f), and these were expanded by another 2,900 models using elastic network modelling, which sampled backbone and loop conformations. Against each of the 3,307 models, we computationally docked the active benzodiazepines, more than 440 inactive compounds from the National Clinical Collection (NCC; http://nihsmr.evotec.com/evotec/sets) library, and 176 property-matched decoy molecules[26]. In each model, five candidate allosteric sites were docked against (Extended Data Fig. 2g), based on the binding regions of aminergic GPCRs, the peptide and antagonist sites of CXCR4, and the muscarinic receptor allosteric site. Iterative cycles of modelling and optimization (Fig. 2b–e) attempted to capture two aspects of ligand binding. First, the activity of

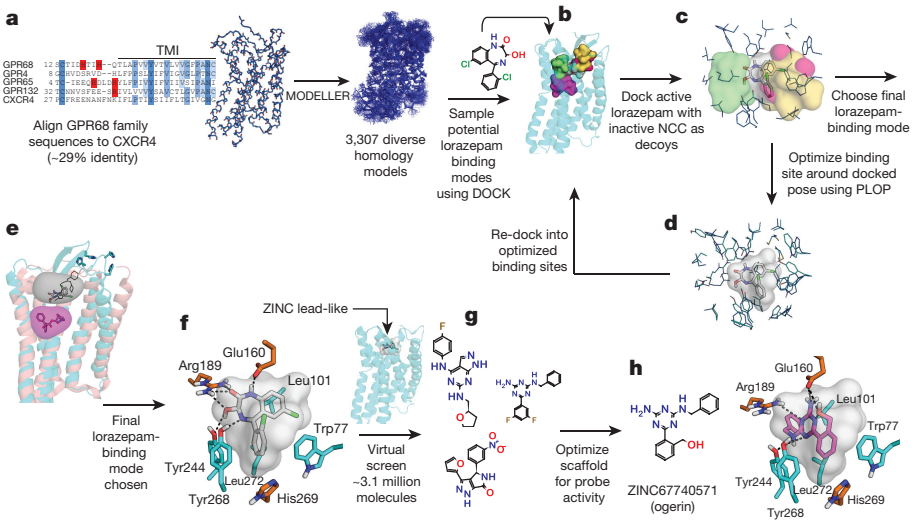


Figure 2 | Virtual screening workflow and predicted location of GPR68 allosteric site. **a**, Sequence alignment of GPR68, GPR4, GPR65 and GPR132 to CXCR4 (details in Extended Data Fig. 2e). **b**, Docking of lorazepam and NCC library to five distinct binding sites (details in Extended Data Fig. 2f). **c**, Models evaluated by their favourable ranking of lorazepam versus decoy molecules. **d**, Optimizing the most favourable lorazepam binding mode. **e**, Optimized lorazepam orientation (grey stick) in GPR68 (cyan ribbon) and M_2 muscarinic receptor (salmon ribbon; Protein Data Bank (PDB) code 4MQT) with allosteric site (grey) and orthosteric site (quinuclidinyl benzilate, magenta). **f**, Lorazepam in its predicted orientation and interactions. **g**, Virtual screen of ZINC subset (~3.1 million molecules) to identify predicted hits. **h**, ZINC67740571 (magenta stick) in its predicted orientation and interactions in GPR68.

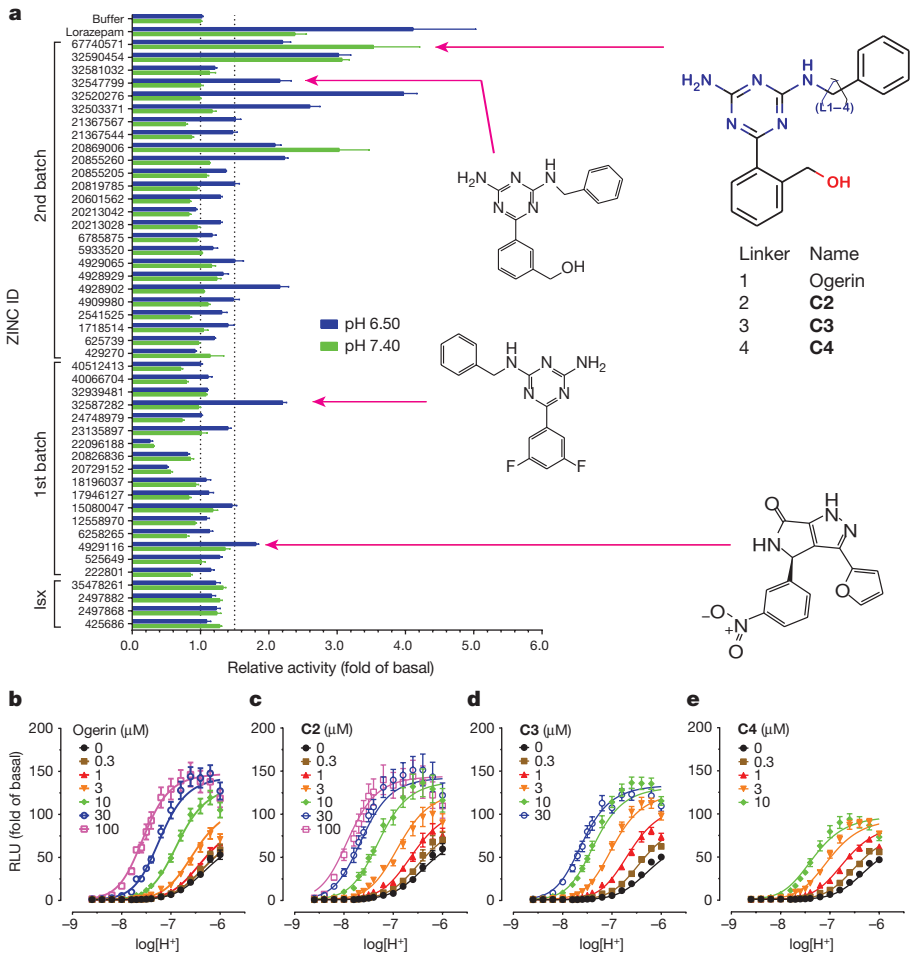


Figure 3 | Identification, characterization and optimization of GPR68-positive allosteric modulators. a, Normalized results of GPR68-mediated cAMP production for selected compounds (ZINC database numbers) are shown; data represent mean ± s.e.m. ($n = 4$–34 measurements) at 10 μM for pH 7.40 and 6.50. Compounds were grouped into a first batch from the first round of virtual docking, and a second batch from the second round of docking. Compounds labelled Isx are isoxazole analogues. Lead compounds ZINC32587282, ZINC4929116, ZINC67740571 (ogerin), its isomer (ZINC32547799) and analogues (**C2**, **C3** and **C4**) with different lengths of linkers, are highlighted. **b**–**e**, Concentration–response curves of normalized data (mean ± s.e.m.; $n = 4$) for ogerin (**b**), **C2** (**c**), **C3** (**d**) and **C4** (**e**) are shown to illustrate the allosteric potentiation of proton and analysed using a standard operational allosteric model. Allosteric parameters are summarized in Supplementary Table 8, and curve-fitting details are in Methods.

the benzodiazepines as PAMs, and second, the role of histidine residues 17, 84, 169 and 269, which are thought to interact with one another in the inactive state, and move apart on protonation at lower pH values[7]. This cycle converged to a stable lorazepam docking pose (Fig. 2f), and to its ranking first among the 622 decoy molecules. This strategy resembles previous ligand-guided docking[27–29], although here the binding site was unknown. In its docked geometry, lorazepam hydrogen bonds with Glu160, Arg189, Tyr244 and Tyr268, and forms non-polar contacts with Trp77, Leu101, Phe173, His269 and Leu272 (Fig. 2f).

To test the modelled lorazepam site, we mutated the Glu160, Arg189 and His269 residues lining the site (Fig. 2f and Extended Data Fig. 2e, f), and determined their roles in proton-mediated cAMP production and calcium release (Extended Data Fig. 4). The His269Phe mutant right-shifted proton concentration–response curves in both assays[7], while substitutions at Arg189 selectively abolished cAMP production. Different substitutions at Glu160 had varying effects at downstream signalling pathways—Glu160Ala left-shifted the proton concentration–response curve and reduced cAMP production, but was inactive in calcium release, while the Glu160Lys and Glu160Gln mutations had modest effects in both pathways (the mutants had little effect on expression, Extended Data Fig. 4c). These substantial and differential effects on downstream coupling support a role for these residues in the functions of GPR68, and are consistent with the modelled binding site for lorazepam.

Seeking optimized PAMs, we computationally docked 3.1 million available lead-like molecules against the putative lorazepam site in GPR68. Overall, more than 3.3 trillion complexes were calculated and scored. From among the top 0.1% of the docking-ranked molecules, 17 were purchased for testing; along with their high docking ranks, these compounds recapitulated key interactions made by lorazepam in its docked model, were chemically diverse and had high-scoring analogues (Supplementary Table 2).

Four of the docking hits increased cAMP production by about 1.5-fold over basal at pH 6.5 (Fig. 3a). Although none was as active as lorazepam, two compounds, ZINC4929116 and ZINC32587282, had hundreds of available analogues. These were docked against the GPR68 model, and 25 were chosen for testing (Fig. 3a and Supplementary Table 3). Thirteen had greater activity than lorazepam, and their pH-dependent potentiation activity clearly indicates allostery. Although dissimilar, lorazepam and ZINC67740571 dock to form many of the same interactions, with the addition of a new predicted hydrogen-bond to Glu160 from the hydroxyl of ZINC67740571 (Fig. 2f, h and Extended Data Fig. 2h).

Ogerin as a selective GPR68 PAM

Ten selected compounds were studied further in functional assays. According to the standard allosteric operational model[30], all were

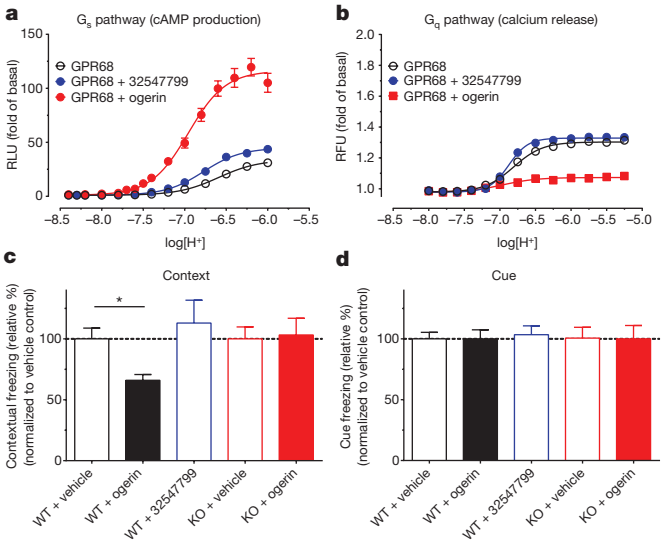


Figure 4 | Ogerin modulates signalling and memory. a, **b**, Ogerin and ZINC32547799 (10 μM) modulate proton-mediated cAMP production (**a**, $n = 4$) and calcium mobilization (**b**, $n = 5$). Data in **a** and **b** are mean ± s.e.m. RFU, relative fluorescence units. **c**, **d**, Ogerin but not its isomer (ZINC32547799) decreased contextual memory retrieval in wild-type (WT; $n = 7$) but not GPR68-knockout (KO; $n = 8$) C57BL/6J male mice (**c**, $F_{(1,27)} = 4.71$, $P < 0.05$ for drug × genotype effect, $P < 0.05$ for ogerin at wild-type mice, two-way analysis of variance (ANOVA), Bonferroni's post-hoc test); both had no effect on cued memory retrieval in either wild-type ($n = 6$) or knockout ($n = 7$) C57BL/6J male mice (**d**). Results (**c**, **d**) were normalized to vehicle control; see also Extended Data Fig. 8d–i.

GPR68 PAMs, lacking intrinsic activity but increasing agonist potency (α-factor) for cAMP production by 1.9–8.2-fold, and increasing efficacy (β-factor) by 1.1–5.6-fold (Supplementary Table 5). It is this ability to shift concentration–response curves leftward and upward (Extended Data Fig. 4b) that are the key characteristics of a PAM. ZINC67740571 had a much higher allosteric effect than lorazepam (Fig. 3b versus Fig. 1d, and Supplementary Table 8); we denoted it 'ogerin' (for OGR1 ligand).

Ogerin and ZINC32547799 are close analogues (Fig. 3a), but each had distinct functional activities (Fig. 4a and Extended Data Fig. 4f, g) and docking poses (Fig. 2 and Extended Data Fig. 2h). Thus, the ortho-hydroxylmethyl group, which differentiates them, may have a key role in determining PAM activity, perhaps because of its ability to hydrogen-bond with Glu160, which the meta-positioned hydroxylmethyl in ZINC32547799 cannot reach. The structure-guided mutants His269Phe and Arg189Leu responded to ogerin and ZINC32547799 differently (Fig. 4a, Extended Data Fig. 4f, g and Supplementary Table 6), supporting the modelled interactions with these residues. Notably, rather than activating, ogerin inhibited proton-mediated calcium release—a pathway-specific function rescued in Arg189Leu and His269Phe (Fig. 4b, Extended Data Fig. 4h, i and Supplementary Table 7). Meanwhile, ZINC32547799 had little effect on calcium release. To determine whether fast kinetics affect the difference between cAMP measurement (under equilibrium) and calcium release (non-equilibrium), we also conducted phosphatidylinositol hydrolysis assays under equilibrium. Ogerin slightly potentiated proton activity here (Extended Data Fig. 4j, k), whereas ZINC32547799 did not. Furthermore, ogerin had minimal PAM activity at the related proton-sensing GPCRs, GPR4 and GPR65 (Extended Data Fig. 2c, d). Ogerin seems to be a functionally selective GPR68 PAM for the agonist proton.

If the ogerin–GPR68 model is relevant, we should be able to leverage it for optimization. We designed a virtual library of more than 600 ogerin analogues and docked each into the GPR68 model (Extended Data Fig. 2h, i). Thirteen high-scoring analogues were synthesized, and three were more active than ogerin (Supplementary Table 9 and Extended Data Fig. 6), including the first and seventh ranked compounds, the latter of which, **C2**, had the greatest allosteric effect, shifting the proton response threefold further to the left than does ogerin, for an α-factor of 22 (Fig. 3a–c and Supplementary Table 8). **C2** differs from ogerin by the addition of a methylene to the benzylamine side chain, which places the phenyl ring deeper into a modelled apolar pocket (Extended Data Fig. 2i). The addition of one or two further methylenes in compounds **C3** and **C4** (Fig. 3a), conversely, reduced allostery (Supplementary Table 8 and Extended Data Fig. 6f), consistent with reduced complementarity to the apolar pocket in the modelled complex.

To investigate ogerin specificity for GPR68 over unrelated targets, which might affect its usefulness as a biological probe, we first computationally screened ogerin and its analogues for off-targets using the Similarity Ensemble Approach (SEA) program[31] against a panel of 2,800 targets. These calculations revealed similarity between the GPR68 ligands and those of only three other GPCRs: the ghrelin and adenosine A_1 and A_{2A} receptors. Subsequent physical profiling against 58 GPCRs, ion channels and transporters (Extended Data Fig. 3) revealed that ogerin had moderate affinity at two GPCRs, 5-hydroxytryptamine 2B (5-HT_{2B}) and the A_{2A} receptor (Extended Data Fig. 5h, i), the latter consistent with the SEA prediction.

Intrigued by the association between the GPR68 PAMs and adenosine receptor antagonists, we computationally screened a library (http://www.tocris.com/dispprod.php?ItemId=5386#.U_s5ZMVdUrU) of 1,120 reagents and drugs against the GPR68 ligands, again using SEA. SLV320, a selective adenosine A_1 antagonist[32], was predicted to be a GPR68 PAM and confirmed by a physical screen of the full library (SLV320 $\alpha\beta = 2.8$) (Extended Data Fig. 7 and Supplementary Table 8), as was a second adenosine receptor antagonist, CGH2466 ($\alpha\beta = 2.9$), and tracazolate ($\alpha\beta = 3.4$), a GABAergic (GABA-mediated) drug that also antagonizes adenosine receptors[33]. Although CGH2466 has the lowest apparent binding constant (K_B) of any GPR68 PAM (48 nM), its allostery is much lower than that of ogerin; additionally, like SLV320 and tracazolate, CGH2466 is a potent phosphodieseterase inhibitor (Extended Data Fig. 7) and had minimal activity in the presence of Ro 20-1724. This previously unknown cross-talk among the GPR68, adenosine and GABA receptor ligands (Extended Data Fig. 7d), along with their activities at phosphodiesterases, should be considered when evaluating the pharmacology of what have been considered specific probes and drugs.

Ogerin as a GPR68 probe

Given its activity and specificity, we sought to explore the downstream signalling and *in vivo* activity of ogerin. In GPR68-expressing HEK293 cells, we found that both ogerin and lorazepam activate the protein kinase A (PKA) and mitogen-activated protein (MAP) kinase pathways (Extended Data Fig. 8a), mimicking the low pH-induced signalling observed with GPR68 receptors in human airway smooth muscle cells[19]. The activation of GPR68 in smooth muscle cells by extracellular acidification is linked to several downstream pathways and biological responses[18,19,22,34–37], which a selective allosteric modulator, such as ogerin, may help to disentangle.

To investigate effects in behaviour associated with modulation of the hippocampus, where GPR68 is highly expressed[17], we evaluated GPR68-knockout and wild-type mice in a learning and memory test, fear conditioning, in which the hippocampus has important roles (Extended Data Figs 8 and 11). In wild-type mice, ogerin attenuated contextual-based fear memory without effects on cue-based memory (Fig. 4c, d). The magnitude of these effects is comparable to those of compounds targeting other hippocampus-expressed GPCRs[38,39], and larger effects are rarely observed without surgical lesion of the hippocampus[40]. Crucially, the administration of ogerin had no effect on memory retrieval in GPR68-knockout mice (Fig. 4c, d), indicating that the *in vivo* effects of ogerin are GPR68-dependent. Furthermore, the less active ogerin isomer, ZINC32547799, had no measurable effect

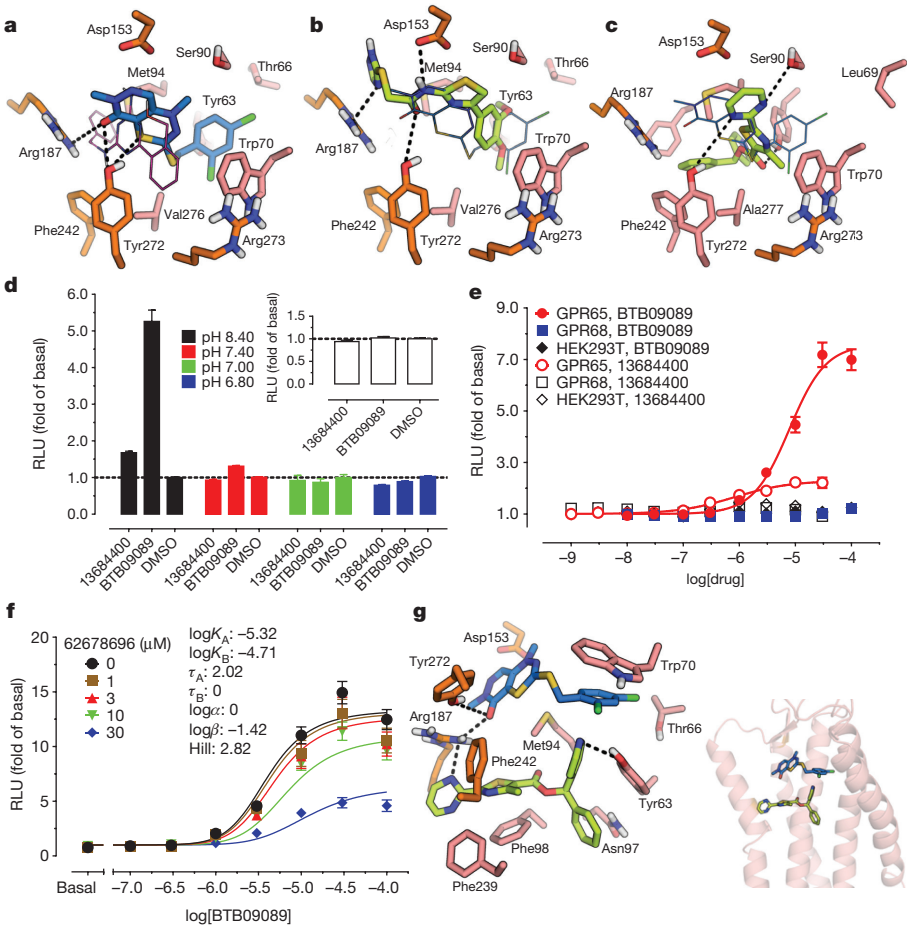


Figure 5 | Discovery of GPR65 allosteric agonist and negative allosteric modulator. a–c, Predicted interactions of BTB09089 (**a**), ZINC13684400 (**b**) and ZINC62678696 (**c**) with GPR65. Overlaid ogerin (thin magenta lines) (**a**) or BTB09089 (thin blue lines) docking poses with GPR68 or GPR65, respectively (**b**, **c**). **d**, ZINC13684400 (30 μM) displayed GPR65 allosteric agonist activity at pH 8.40 but not at lower pH or in control cells ($n = 32$ measurements). **e**, ZINC13684400 as a GPR65 agonist at pH 8.40 ($n = 3$). **f**, ZINC62678696 shifts BTB09089 curves downward at pH 8.40 ($n = 4$). K_A and K_B are the equilibrium binding affinities of the orthosteric agonist proton (A) and allosteric modulator (B), respectively. Normalized results (**d**–**f**) are mean ± s.e.m., and curves were analysed using a four-parameter logistic function (**e**) or a standard operational allosteric model (**f**). **g**, Predicted ternary complex between GPR656, ZINC62678696 and BTB09089, detailed interactions (left) and overall orientation in the GPR65 structure (right).

on learning and memory in wild-type mice (Fig. 4c, d and Extended Data Fig. 8d–i). The effects of ogerin thus support a role for GPR68 in hippocampal-associated memory.

General applicability of the approach

To explore the broader usefulness of this approach, we sought ligands for GPR65, another understudied pH-sensing receptor, which shares 37% sequence identity to GPR68. We found that a recently reported GPR65 agonist BTB09089 (ref. 41) is an allosteric agonist of GPR65 (Fig. 5d, e and Extended Data Fig. 10a). We used BTB09089 to anchor modelling of GPR65, generating 500 homology models templated on GPR68. The final docked GPR65–BTB09089 model resembles that of GPR68–ogerin, with several side-chain substitutions in the putative binding site (Fig. 5a).

We docked the same 3.1-million compounds against the GPR65 model, purchasing 45 new molecules for testing (Fig. 5a–c and Supplementary Table 10). ZINC13684400 showed agonist activity of more than twofold of basal at GPR65, with a potency of 500 nM, without measurable activity at control cells (Fig. 5e and Extended Data Fig. 9). As with BTB09089, ZINC13684400 did not potentiate proton efficacy at GPR65 (Fig. 5d), but acted as an allosteric agonist. To test the model, three residues modelled to interact with both BTB09089 and ZINC13684400, Arg187, Phe242 and Tyr272, were mutated, as was Asp153, which appears to only hydrogen-bond with ZINC13684400 (Fig. 5b). Arg187Leu, Phe242Ala and Tyr272Ala reduced the activity of both compounds (Extended Data Fig. 10f, g), whereas Asp153Ala had no effect on BTB09089 but much reduced the activity of ZINC13684400, consistent with the model. Several other docking hits inhibited GPR65 when the receptors were activated by protons or by BTB09089, including ZINC62678696 (Extended Data Fig. 10b–d). Unexpectedly, ZINC62678696 does not compete with BTB09089, as predicted, but rather acts as a BTB09089 negative allosteric modulator (Fig. 5f), suggesting that the two molecules can bind to GPR65 simultaneously (Fig. 5g).

Discussion

A combined empirical and structure-based approach discovered potent PAMs at the understudied receptor GPR68, and an allosteric agonist and negative allosteric modulators for the understudied GPR65. This supports the usefulness of the approach for illuminating the 'dark matter' of the GPCRs—the 38% of non-olfactory GPCR targets whose ligands and function are understudied or unknown[1]. Whereas truly high-throughput screens are impractical for targets of unknown function, lower-throughput screens are often feasible. Although the hits from such a screen may be unsuitable as probes, they can anchor computational screens for more optimized compounds. Correspondingly, we would not ordinarily expect docking to succeed against models of a target that shares only 29% sequence identity with its nearest template. By calculating several thousand models, and insisting that the relevant ones are those that prioritize active over inactive molecules,

functionally relevant models are prioritized. The new ligands that emerged are specific for the target and one is active *in vivo*, supporting their use as chemical probe for the function of GPR68.

Pharmacologically, the most unexpected observation was the activity of GPR68 in learning and memory. Previous studies in GPR68-knockout mice revealed only modest phenotypic changes[16,21,22], none in higher brain function, even though GPR68 is most highly expressed in the brain. Ogerin transiently and reversibly reduced contextual-based fear memory in wild-type but not GPR68-knockout mice, consistent with on-target activity *in vivo*. In hindsight, this is perhaps only accessible to chemical modulators, which can have PAM activities. Inhibitory genetic perturbations, such as knockouts or knockdowns, although crucial to demonstrating on-target activity through chemical genetic epistasis, cannot on their own reveal such activation-based modulation.

Deorphanizing a receptor can also illuminate its off-target roles for known drugs. The observation that lorazepam and its primary metabolite, desmethyldiazepam, are GPR68 PAMs may clarify several of the idiosyncratic effects of this widely used anxiolytic. Lorazepam, uniquely among benzodiazepines, can treat catatonia, an effect proposed to involve an unknown secondary target[42]. GPR68 may have a role in this efficacy, as both drug and metabolite reach micromolar concentrations in plasma during treatment[43].

Certain caveats bear airing. The combination of empirical and computational screens will not work for all orphan receptors. GPCRs that are poorly expressed or non-functional in yeast or transfected cells will be problematic, and some orphans will simply not recognize any of the molecules screened in the small empirical libraries. Also, some orphans will bear too little similarity to templates of known structure to support accurate modelling. Even those that do work will demand cycles of testing and optimization, which was crucial for both GPR65 and GPR68.

These cautions should not obscure the key observations from this study—that combining empirical and structure-based screening led to a probe molecule that reveals some of the functions of GPR68. The finding that ogerin potentiates GPR68 activation and downstream MAP kinase pathways, and previous observations that the receptor mediates airway inflammation, enables campaigns for GPR68 PAMs that may regulate respiratory inflammatory responses. Uniquely as PAMs, these compounds would have fidelity to the natural spatial and temporal activation of GPR68. Correspondingly, the role of GPR68 in anxiety offers a new route to treating this condition and related CNS disorders, an area in need of new therapeutic modalities[44]. Methodologically, this approach may have broad application to illuminating the function of the dark matter of the genome, that still large area of pharmacology in which targets are known, but function is hidden.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 2 January; accepted 4 September 2015.
Published online 9 November 2015.

1. Roth, B. L. & Kroeze, W. K. Integrated approaches for genome-wide interrogation of the druggable non-olfactory G protein-coupled receptor superfamily. *J. Biol. Chem.* **290,** 19471–19477 (2015).
2. Davenport, A. P. *et al.* International Union of Basic and Clinical Pharmacology. LXXXVIII. G protein-coupled receptor list: recommendations for new pairings with cognate ligands. *Pharmacol. Rev.* **65,** 967–986 (2013).
3. Chung, S., Funakoshi, T. & Civelli, O. Orphan GPCR research. *Br. J. Pharmacol.* **153** (suppl. 1), S339–S346 (2008).
4. Knapp, S. *et al.* A public–private partnership to unlock the untargeted kinome. *Nature Chem. Biol.* **9,** 3–6 (2013).
5. Ferguson, F. M. *et al.* Targeting low-druggability bromodomains: fragment based screening and inhibitor design against the BAZ2B bromodomain. *J. Med. Chem.* **56,** 10183–10187 (2013).
6. Leung, D., Hardouin, C., Boger, D. L. & Cravatt, B. F. Discovering potent and selective reversible inhibitors of enzymes in complex proteomes. *Nature Biotechnol.* **21,** 687–691 (2003).
7. Ludwig, M. G. *et al.* Proton-sensing G-protein-coupled receptors. *Nature* **425,** 93–98 (2003).
8. Mogi, C. *et al.* Sphingosylphosphorylcholine antagonizes proton-sensing ovarian cancer G-protein-coupled receptor 1 (OGR1)-mediated inositol phosphate production and cAMP accumulation. *J. Pharmacol. Sci.* **99,** 160–167 (2005).
9. Li, J. *et al.* Ovarian cancer G protein coupled receptor 1 suppresses cell migration of MCF7 breast cancer cells via a Gα12/13-Rho-Rac1 pathway. *J. Mol. Signal.* **8,** 6 (2013).
10. Singh, L. S. *et al.* Ovarian cancer G protein-coupled receptor 1, a new metastasis suppressor gene in prostate cancer. *J. Natl. Cancer Inst.* **99,** 1313–1327 (2007).
11. Schneider, J. W. *et al.* Coupling hippocampal neurogenesis to brain pH through proneurogenic small molecules that regulate proton sensing G protein-coupled receptors. *ACS Chem. Neurosci*. **3,** 557–568 (2012).
12. Frick, K. K., Krieger, N. S., Nehrke, K. & Bushinsky, D. A. Metabolic acidosis increases intracellular calcium in bone cells through activation of the proton receptor OGR1. *J. Bone Miner. Res.* **24,** 305–313 (2009).
13. Komarova, S. V., Pereverzev, A., Shum, J. W., Sims, S. M. & Dixon, S. J. Convergent signaling by acidosis and receptor activator of NF-κB ligand (RANKL) on the calcium/calcineurin/NFAT pathway in osteoclasts. *Proc. Natl Acad. Sci. USA* **102,** 2643–2648 (2005).
14. Yang, M. *et al.* Expression of and role for ovarian cancer G-protein-coupled receptor 1 (OGR1) during osteoclastogenesis. *J. Biol. Chem.* **281,** 23598–23605 (2006).
15. Russell, J. L. *et al.* Regulated expression of pH sensing G protein-coupled receptor-68 identified through chemical biology defines a new drug target for ischemic heart disease. *ACS Chem. Biol.* **7,** 1077–1083 (2012).
16. Mohebbi, N. *et al.* The proton-activated G protein coupled receptor OGR1 acutely regulates the activity of epithelial proton transport proteins. *Cell. Physiol. Biochem.* **29,** 313–324 (2012).
17. Regard, J. B., Sato, I. T. & Coughlin, S. R. Anatomical profiling of G protein-coupled receptor expression. *Cell* **135,** 561–571 (2008).
18. Chen, Y. J., Huang, C. W., Lin, C. S., Chang, W. H. & Sun, W. H. Expression and function of proton-sensing G-protein-coupled receptors in inflammatory pain. *Mol. Pain* **5,** 39 (2009).
19. Saxena, H. *et al.* The GPCR OGR1 (GPR68) mediates diverse signalling and contraction of airway smooth muscle in response to small reductions in extracellular pH. *Br. J. Pharmacol.* **166,** 981–990 (2012).
20. Wang, J., Sun, Y., Tomura, H. & Okajima, F. Ovarian cancer G-protein-coupled receptor 1 induces the expression of the pain mediator prostaglandin E2 in response to an acidic extracellular environment in human osteoblast-like cells. *Int. J. Biochem. Cell Biol.* **44,** 1937–1941 (2012).
21. Li, H. *et al.* Abnormalities in osteoclastogenesis and decreased tumorigenesis in mice deficient for ovarian cancer G protein-coupled receptor 1. *PLoS ONE* **4,** e5705 (2009).
22. Aoki, H. *et al.* Proton-sensing ovarian cancer g protein-coupled receptor 1 on dendritic cells is required for airway responses in a murine asthma model. *PLoS ONE* **8,** e79985 (2013).
23. Mogi, C., Nakakura, T. & Okajima, F. Role of extracellular proton-sensing OGR1 in regulation of insulin secretion and pancreatic β-cell functions. *Endocr. J.* **61,** 101–110 (2013).
24. Okajima, F. Regulation of inflammation by extracellular acidification and proton-sensing GPCRs. *Cell. Signal.* **25,** 2263–2271 (2013).
25. Dong, S., Rogan, S. C. & Roth, B. L. Directed molecular evolution of DREADDs: a generic approach to creating next-generation RASSLs. *Nature Protocols* **5,** 561–573 (2010).
26. Mysinger, M. M. & Shoichet, B. K. Rapid context-dependent ligand desolvation in molecular docking. *J. Chem. Inf. Model.* **50,** 1561–1573 (2010).
27. Evers, A. & Klebe, G. Ligand-supported homology modeling of G-protein-coupled receptor sites: models sufficient for successful virtual screening. *Angew. Chem. Int. Ed. Engl.* **43,** 248–251 (2004).
28. Cavasotto, C. N. *et al.* Discovery of novel chemotypes to a G-protein-coupled receptor through ligand-steered homology modeling and structure-based virtual screening. *J. Med. Chem.* **51,** 581–588 (2008).
29. Katritch, V., Rueda, M., Lam, P. C.-H., Yeager, M. & Abagyan, R. GPCR 3D homology models for ligand screening: lessons learned from blind predictions of adenosine A2a receptor complex. *Proteins* **78,** 197–211 (2010).
30. Leach, K., Sexton, P. M. & Christopoulos, A. Allosteric GPCR modulators: taking advantage of permissive receptor pharmacology. *Trends Pharmacol. Sci.* **28,** 382–389 (2007).
31. Keiser, M. J. *et al.* Relating protein pharmacology by ligand chemistry. *Nature Biotechnol.* **25,** 197–206 (2007).
32. Kalk, P. *et al.* The adenosine A1 receptor antagonist SLV320 reduces myocardial fibrosis in rats with 5/6 nephrectomy without affecting blood pressure. *Br. J. Pharmacol.* **151,** 1025–1032 (2007).
33. Thompson, S.-A., Wingrove, P. B., Connelly, L., Whiting, P. J. & Wafford, K. A. Tracazolate reveals a novel type of allosteric interaction with recombinant γ-aminobutyric $acid_A$ receptors. *Mol. Pharmacol.* **61,** 861–869 (2002).
34. Tomura, H. *et al.* Prostaglandin I_2 production and cAMP accumulation in response to acidic extracellular pH through OGR1 in human aortic smooth muscle cells. *J. Biol. Chem.* **280,** 34458–34464 (2005).
35. Ichimonji, I. *et al.* Extracellular acidification stimulates IL-6 production and Ca^{2+} mobilization through proton-sensing OGR1 receptors in human airway smooth muscle cells. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **299,** L567–L577 (2010).

36. Liu, J. P. *et al.* Ovarian cancer G protein-coupled receptor 1-dependent and -independent vascular actions to acidic pH in human aortic smooth muscle cells. *Am. J. Physiol. Heart Circ. Physiol.* **299,** H731–H742 (2010).
37. Matsuzaki, S. *et al.* Extracellular acidification induces connective tissue growth factor production through proton-sensing receptor OGR1 in human airway smooth muscle cells. *Biochem. Biophys. Res. Commun.* **413,** 499–503 (2011).
38. Gravius, A., Barberi, C., Schäfer, D., Schmidt, W. J. & Danysz, W. The role of group I metabotropic glutamate receptors in acquisition and expression of contextual and auditory fear conditioning in rats – a comparison. *Neuropharmacology* **51,** 1146–1155 (2006).
39. Daumas, S. *et al.* Transient activation of the CA3 Kappa opioid system in the dorsal hippocampus modulates complex memory processing in mice. *Neurobiol. Learn. Mem.* **88,** 94–103 (2007).
40. Phillips, R. G. & LeDoux, J. E. Differential contribution of amygdala and hippocampus to cued and contextual fear conditioning. *Behav. Neurosci.* **106,** 274–285 (1992).
41. Onozawa, Y. *et al.* Activation of T cell death-associated gene 8 regulates the cytokine production of T cells and macrophages *in vitro*. *Eur. J. Pharmacol.* **683,** 325–331 (2012).
42. Pompéia, S., Manzano, G. M., Tufik, S. & Bueno, O. F. What makes lorazepam different from other benzodiazepines? *J. Physiol. (Lond.)* **569,** 709 (2005).
43. Greenblatt, D. J. *et al.* Clinical pharmacokinetics of lorazepam. I. Absorption and disposition of oral 14C-lorazepam. *Clin. Pharmacol. Ther.* **20,** 329–341 (1976).
44. Schoepp, D. D. Where will new neuroscience therapies come from? *Nature Rev. Drug Discov.* **10,** 715–716 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported by National Institutes of Health (NIH) grants U01104974 (B.L.R., B.K.S. and W.K.K.), R01 DA017204 (B.L.R. and W.K.K.) and the National Institute of Mental Health Psychoactive Drug Screening Program (NIMH PDSP) (X.-P.H., H.Z., M.S.F., W.K.K., T.J.M., A.J. and B.L.R.), the Michael Hooker Chair for Protein Therapeutics and Translational Proteomics to B.L.R.; Genentech Foundation Predoctoral Fellowship (J.K.); NIH grants GM59957 and GM71896 (B.K.S.) and the Structural Genomics Consortium (B.K.S.); grant P01 HL114471 (R.B.P. and D.A.D.); NICHD grant U54 HD079124 (M.S.M., K.A.S., V.N.); NIH grant U19MH082441 (B.L.R., J.J. and X.C.). We thank Mark Pausch (Merck & Co.) for providing us G_s- and G_q-yeast strains for yeast screening assays.

Author Contributions X.-P.H. subcloned GPR68 for yeast screening, made GPR68 and GPR65 mutants, designed, carried out cell-based screening assays, analysed results, and wrote the paper. J.K. designed and developed homology models, carried out docking screens, analysed results, and wrote the paper. W.K.K. set up and performed yeast screening assays, analysed results, and wrote the paper. H.Z. and M.S.F. designed, performed *in vivo* fear-conditioning studies, analysed results, and wrote the paper. M.S.F. and B.L.R. dubbed ZINC67740571 'ogerin'. B.H.K. created the GPR68-knockout mice. S.S.M., K.A.S. and V.N. carried out initial phenotypic characterization, analysed results, and wrote the paper. X.C. and J.J. synthesized ZINC32547799, ZINC67740571 (ogerin) and ogerin analogues (compounds 33548–33561, **C3** and **C4**) for functional assays and *in vivo* studies, and wrote the paper. T.J.M. carried out radioligand binding assays. A.J. prepared drug plates and plasmids for initial screening. R.B.P. and D.A.D. designed and carried out anti-haemagglutinin immunoblot assays, analysed results, and wrote the paper. S.W. designed primers, prepared Flag-tagged GPR68 wild-type and mutant plasmids, performed anti-Flag western blot assays, and analysed results. T.K. analysed results and wrote the paper. B.L.R. and B.K.S. coordinated and supervised the project, and with the other authors wrote the paper.

Additional Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to B.L.R. (bryan_roth@med.unc.edu) or B.K.S. (bshoichet@gmail.com).

METHODS

Chemicals, reagents and cells lines. Chemicals and reagents used in this study, if not specified otherwise, were purchased from commercial sources (Sigma, Tocris, Fisher Scientific, or specified in Supplementary Tables 2 and 3 of chemical structures) or synthesized as outlined in the Supplementary Information. HEK293 (ATCC CRL-1573; 60113019; certified mycoplasma free and authentic by ATCC) and HEK293-T (HEK293T; ATCC CRL-11268; 59587035; certified mycoplasma free and authentic by ATCC) cells were from the ATCC. Cells were also validated by analysis of short tandom repeat (STR) DNA profiles and these profiles showed 100% match at the STR database from ATCC. Ogerin and its inactive analogue ZINC32547799 are available for use as chemical probes from Sigma-Aldrich (ogerin: SML1482, ZINC32547799: SML1483).

Homology modelling. The alignment for the construction of the GPR68 models was generated using PROMALS3D, and homology models were built with MODELLER-9v8 (ref. 45), using the crystal structure of the chemokine CXCR4 receptor (PDB code 3ODU) as the template (Extended Data Fig. 2f). This alignment was also used to generate 500 models of GPR65 directly from the final GPR68 model. The initial alignment included both human and mouse sequences of GPR68, as well as those of its closest homologue, GPR4. These were aligned against the whole human C-X-C chemokine receptor family. The alignment was manually edited to: remove the amino and carboxy termini that extended past the template structure, remove the engineered T4 lysozyme, and create different alignments of the flexible and non-conserved second extracellular loop (the final result is given in the provided alignment, Extended Data Fig. 2f). A total of 407 models were built directly based on the CXCR4 crystal structure, using MODELER-9v8 (ref. 45), while five more were built from each of 580 elastic network models (ENMs), produced by the program 3K-ENM[46], for a total of 3,307 models built during each iterative round of model refinement. Models with constraints between pairs of extracellular His residues (His17–His169, His17–His269, His17–His84 and His84–His169) to mimic the inactive state of the protein were generated by enforcing a distance constraint of 2.7 Å between the imidazole nitrogens, with a standard deviation of 0.1 Å. Confirmed active compounds and analogues using CXCR4-based model had neither agonist nor antagonist activity at CXCR4 receptors (Extended Data Fig. 5j, k).

Model evaluation. Before docking, the second extracellular loop (EL2), between residues 161–177, was removed from each GPR68 model. Models were ranked on the basis of prioritizing active benzodiazepines (lorazepam and desmethyldiazepam) over the rest of the inactive NCC library that was used in the yeast screen, as well as over property-matched decoys. In addition, the docked pose of lorazepam had to form a hydrogen bond from its N–H group to a polar side chain in GPR68. Five different sites were sampled for possible lorazepam binding, based on the locations of the co-crystallized CXCR4 small molecule antagonist 1T1t (in PDB code 3ODU), cyclic peptide CVX15 (in PDB code 3OE0), and the positions of the biogenic amines crystallized with the β_2-adrenergic receptor (PDB code 2RH1) and the dopamine D_3 receptor (PDB code 3PBL). The entire NCC library was docked to each of the five sub-sites for several rounds of iterative binding site refinement. In each round, the top-ranked models were examined for a binding pose that made hydrophobic and electrostatic interactions with the receptor, including the key N–H hydrogen bond. Residues within 6 Å of the lorazepam pose were minimized around the docked ligand with PLOP[47]. The NCC library was then re-docked into this optimized binding site for each model. This refinement continued for several cycles until the top-ranked models all converged to the same lorazepam pose. Once the final model was chosen, we built the EL2 back onto the receptor using MODELLER-9v8 (ref. 45) and optimized 1,000 different EL2 conformations around the lorazepam pose with PLOP. Finally, we docked the NCC library back into these 1,000 different EL2-GPR68 structures, and chose a final model that retained the previous pose and prioritized the active over the inactive compounds. The GPR65 model was generated similarly, using the pose of BTB09089 as the primary selection criterion, although in this case the EL2 was always present. To determine the ternary complex model of ZINC62678696 and BTB09089, ZINC62678696 was docked to the putative binding site in the GPR65 model with BTB09089 present. Then, both ligands were minimized with PLOP. Next, the side chains of the GPR65 binding pocket were allowed to relax, and, finally, BTB09089 and ZINC62678696 were simultaneously minimized again with PLOP. Structural models (PDB files) of characteristic GPR68-modelled complexes (with ogerin or lorazepam) and GPR65-modelled complexes (with BTB09089 or BTB09089 and ZINC62678696) are shown in the Supplementary Data.

Virtual screens. We used DOCK 3.6 to screen the ZINC database (Results). The flexible ligand sampling algorithm in DOCK 3.6 superimposes atoms of the docked molecule onto binding site matching spheres, which represent favourable positions for individual ligand atoms. Forty-five matching spheres were used, using the previous refinement round's pose of lorazepam. The degree of ligand sampling is determined by the bin size, bin size overlap and distance tolerance, set at 0.4 Å, 0.1 Å and 1.5 Å, respectively, for both the matching spheres and the docked molecules. The complementarity of each ligand pose was scored as the sum of the receptor–ligand electrostatic and van der Waals' interaction energies, and corrected for context-dependent ligand desolvation. Partial charges from the united-atom AMBER force field were used for all receptor atoms; ligand charges and initial solvation energies were calculated using AMSOL[48,49] (http://comp.chem.umn.edu/amsol/). The best-scoring conformation of each docked molecule was then subjected to 100 steps of rigid-body minimization.

Selection of potential ligands for testing. We docked the approximately 3.1 million commercially available molecules of the lead-like subset of the ZINC database to the final GPR68 and GPR65 models. The full hit list was automatically filtered to remove molecules that possess high-internal-energy, non-physical conformations, which are not well-modelled by our scoring function. The reported rankings reflect this filtering. From the top 0.1% (~3,000 molecules) of the docked ranking list, 17 compounds were chosen for testing, based on complementarity to the binding site and presence of predicted electrostatic interactions with Glu160, Arg189, Tyr244, Tyr268 and His269, mimicking those predicted for lorazepam. For GPR65, compounds were chosen based on complementarity to the binding site and similarity to the predicted binding pose of BTB09089, modelled to interact with Asp153, Arg187 and Tyr272, and by aromatic stacking with Trp70.

***In silico* lead profiling.** To examine specificity and to discover other potential GPCR targets for the newly discovered GPR68 PAMs, we used the SEA program[31,50], which compares individual ligands and sets of ligands to the ligand sets for multiple targets; two targets are related, or a particular ligand is predicted to modulate a target, if the sets of ligands are related to one another. Here, the query set was all of the new GPR68 PAMs, which was screened against either the 2,512 ligand–target set with activity of 10 μM or better from the ChEMBL12 database[51], or against the Tocris Mini library.

Receptor constructs and yeast growth assays. Twenty-four human GPCR plasmids (GPR1, GPR4, GPR15, GPR31, GPR39, GPR41, GR43, GPR45, GPR55, GPR57, GPR58, GPR62, GPR65, GPR68, GPR83, GPR84, GPR87, GPR88, GPR123, GPR132, GPR133, GPR157, GPR161 and ADCYAP1R1) were obtained from http://cdna.org, subcloned into the multiple cloning site of the yeast high copy number plasmid p426GPD (ref. 52) and were confirmed by full-length sequencing (Eton Bioscience). The yeast strains used were provided by M. Pausch (Merck) and have been previously described[53] and used by us[25,54] ; MPY578t (G_i yeast), MPY578q5 (G_q yeast) and MPY578s5 (G_s yeast) express chimaeric G proteins in which the last five amino acids of the yeast Gα protein are replaced with their mammalian G_i, G_q or G_s homologues, respectively. These strains contain the *HIS3* gene under the control of the *FUS1* promoter. GPCR transformants in yeast were selected and maintained on synthetic defined (SD) media lacking uracil (Clontech). $GPR68_q$ indicates the GPR68 paired with G_q yeast; while $GPR4_s$ indicates GPR4 paired with G_s yeast, and similarly for the other GPCRs. The yeast screening assays were carried out as described previously[25]. Assays were set up in 96-well flat-bottom clear assay plates that contained 50 μl of test compound at 40 μM (final concentration of 10 μM, in triplicate) diluted in SD-His-Ura medium (Clontech), 50 μl of 3-amino-1,2,4-triazole (3-AT) at 4× concentration diluted in SD-His-Ura medium (pH 5.4), and 100 μl of yeast cell suspension diluted in SD-His-Ura medium to a final $A_{600\,nm}$ of 0.02. Growth was at 30 °C for 2–5 days. Before measurement of cell growth, cells were re-suspended by repeated gentle pipetting to ensure uniform suspension of cells. Cell growth was measured by absorbance at 600 nm in a microplate reader (POLARstar Omega, BMG Biotech). After culling of data from obviously contaminated wells, the $A_{600\,nm}$ values of each individual well were adjusted as follows: 100× ($A_{600\,nm}$ of test well − $A_{600\,nm}$ of plate median value) to give percentage growth stimulation (positive values), or percentage growth inhibition (negative value) in the form of mean ± s.e.m. of three wells.

To measure and control constitutive activity or leaky *HIS* expression, each receptor–yeast combination was plated as above in the absence of ligand over a range of concentrations of 3-AT. Concentrations of 3-AT that showed moderate yeast growth (that is, *A* values of 0.2–0.6) after 2 days at 30 °C were used in assays for drug screening. To measure concentration-dependent activity, various concentrations of cognate ligands diluted in SD-His-Ura medium were incubated with transformed yeast and appropriate concentrations of 3-AT for 2 days at 30 °C.

Site-directed mutagenesis. The GPR68 plasmid was obtained from http://cdna.org. Mutation of Glu160Ala, Glu160Lys, Glu160Gln, Arg189Leu, Arg189Met and His269Phe in the GPR68 and mutation of Asp153Ala, Arg187Leu, Phe242Ala and Tyr272Ala in the GPR65 were introduced with Agilent's QuikChange II site-directed mutagenesis kit and confirmed by sequencing. To tag the receptors for comparing receptor expression levels with immunoblotting, Flag epitope tag was inserted at the C terminus of the GPR68 wild-type and mutant receptors, also using the QuikChange II site-directed mutagenesis kit. Insertion was confirmed by sequencing.

Split-luciferase based cAMP reporter assays with proton receptors. GPR4, GPR65 and GPR68 plasmids were obtained from http://cdna.org. GPR68 mutations were made and confirmed as above. Receptor-mediated G_s activation was measured using a split-luciferase reporter assay (GloSensor cAMP assay, Promega). In brief, HEK293T cells were transiently co-transfected with receptor DNA and the GloSensor cAMP reporter plasmid (GloSensor 7A). Transfected cells were plated in poly-L-Lys-coated 384-well white clear bottom cell culture plates in DMEM supplemented with 1% dialysed FBS at a density of 15,000 cells per well in a total volume of 40 µl for a minimum of 6 h. Before assays, culture medium was removed and cells were incubated with luciferin (4 mM prepared in drug buffer, pH 8.4) for 90 min at 37 °C. The drug buffer was made with 1× HBSS supplemented with 10 mM HEPES and 10 mM MES modified from[19]. TAPS was added to accommodate higher pH values for some assays; no difference was observed between different buffers under the same pH conditions. Cells plated at pH 8.4 for 6 h generated the same H^+ concentration–response curves as those plated at pH 7.4. To make individual pH solutions, the pH was adjusted with NaOH and measured at room temperature with a pH 211 Microprocessor pH meter (Hanna Instruments). To measure modulator activity under different pH conditions, modulator was mixed with pH solutions before adding to cells. To achieve the goal that drug solutions were delivered at the correct pH values, luciferin solution was removed from cell plates before addition of drug solutions at predetermined pH values. To improve solubility for some hydrophobic compounds, 1 mg ml^{-1} BSA was added to drug solutions, and it had no effect on H^+ concentration–response curves. For G_s protein activity (cAMP production), the cell plate was usually incubated at room temperature for 20 min before being counted in a luminescence counter. Results were analysed using GraphPad Prism.

Allosteric operational model and data analysis. To estimate allosteric parameters, results were fitted to the allosteric operational model[30,55] as shown in the following equation:

$$\text{Response} = \text{basal} + (E_{max} - \text{basal}) \times \frac{(\tau_A[A](K_B + \alpha\beta[B]))^n}{([A]K_B + K_AK_B + K_A[B] + \alpha[A][B])^n + (\tau_A[A](K_B + \alpha\beta[B]))^n}$$

In which:

(1) Response is the measured activity in the form of RLUs for measurement of cAMP production. If the results were normalized, the 'response' is RLU in fold of basal (with buffer control as basal).

(2) E_{max} is a system parameter, representing the maximal possible response of the system, and this value was normally constrained to the maximal reading of the corresponding experiment.

(3) Basal is the baseline in the absence of test ligand, and is constrained to the baseline of the corresponding experiment. If results were normalized to fold of basal, the 'basal' was usually 1.0.

(4) [A] and [B] represent concentrations of the orthosteric and allosteric ligands, respectively. In the case of GPR68, A is proton.

(5) K_A and K_B are the equilibrium dissociation constants of the orthosteric agonist proton (A) and allosteric modulator (B), respectively. To facilitate curve-fitting with the model, K_A is usually fixed to the binding affinity determined from traditional radioligand binding assays under the assumption that the experimentally derived binding affinity is not significantly different from the functional affinity under the condition for corresponding functional assay. Since proton binding affinity is not a measurable parameter in this assay system, the proton K_A is therefore constrained to the corresponding proton potency (EC_{50}, the proton concentration for half-maximal response) value in the absence of the allosteric ligand, under the assumption that the proton potency is not significantly different from its binding affinity when the cAMP production assay is carried out. Since protons are present at relevant concentrations at physiological pH values, for a proton receptor K_B is largely a fitting parameter without a clear physical meaning.

(6) The term τ_A is the orthosteric agonist proton efficacy parameter. Since allosteric modulators in this study showed no agonist activity, the allosteric modulator efficacy τ_B is therefore 0 and not included in the function.

(7) The term n is the slope factor linking receptor occupancy to response. Steep slopes in this study indicated high cooperativity between proton binding and receptor activation, probably reflecting the fact that the proton receptors operate within a narrow physiological pH range.

(8) The allosteric parameter α defines the mutual effect between the orthosteric agonist A and the allosteric modulator B ($\alpha > 1$ for increased affinity and $\alpha < 1$ for reduced affinity); while β defines the allosteric effect on agonist efficacy ($\beta > 1$ for increased efficacy and $\beta < 1$ for reduced efficacy).

With K_A, basal and E_{max} constrained to their corresponding values, the parameters K_B, τ_A, α, β and n are globally shared fitting parameters for a family of proton concentration–response curves in the absence and presence of increasing concentrations of a test allosteric modulator. With the above settings, most curves could be easily fitted to generate reasonable parameters. If Prism could not fit the curves, but generated 'ambiguous fitting' results, the α value was then manually constrained to an initial fitting value and systematically changed with small increments or decrements until the highest stable high affinity value (K_B) was reached. For GPR65 and GPR68, K_B represents the allosteric binding affinity in the absence of protons, which is unmeasurable and thus has little physical meaning. The value $K_B/(1+\alpha)$ represents the binding affinity of an allosteric ligand in the presence of protons, which could be estimated experimentally. For convenience, we call $K_B/(1+\alpha)$ the 'Biochemical binding affinity, K_{bB}' (Supplementary Table 8) for an allosteric ligand in the presence of an orthosteric agonist (in this case, H^+).

Calcium mobilization assays. HEK293T cells were transfected and plated into poly-L-Lys-coated 384-well black clear bottom cell culture plates in DMEM supplemented with 1% dialysed FBS, at a density of 15,000 cells in 40 µl per well for overnight. Before the assay, medium was removed and cells were loaded with Fluo-4 Direct calcium dye (Invitrogen) for 60 min at 37 °C in a 5% CO_2 atmosphere. The calcium dye was prepared in drug buffer supplemented with 2.5 mM probenecid, pH 8.0. Proton solutions were made with 1× HBSS, 7 mM HEPES, 7 mM HEPPS and 7 mM MES, and pH was adjusted with NaOH. Drug additions and fluorescence intensity measurement were carried out in a $FLIPR^{TETRA}$, which was programmed to add drug solutions to cells while recording fluorescence intensity. To measure proton concentration–responses, 10 µl of pH pre-determined solutions were added to each well (with 20 µl calcium dye) while fluorescence intensity was recorded during and after addition for 4 min (one reading per second). The addition procedure was configured in such a way (30 µl per second at height of 10 µl above cells) that local proton concentrations for cells were essentially the same as in the pH working solutions at the moment of addition. Fluorescence intensities reached peak values within 30 s after drug addition. To determine the effects of modulators on proton responses, the protocol was modified slightly. In brief, cells were loaded with calcium dye as above, but only at 15 µl per well. The $FLIPR^{TETRA}$ was programmed to first add 5 µl of 4× test compound (final concentration of 10 µM before addition of 10 µl of pH solutions) prepared with the same drug buffer at pH 8.0 (buffer alone served as a control). After a total of 10 min of reading and incubation, 10 µl of the pH solutions were added and the fluorescence intensity was recorded exactly the same way as above. Results (fluorescence intensity in fold of basal) were exported and analysed in GraphPad Prism. For calcium mobilization assays with $5\text{-}HT_{2B}$ receptors, HEK293 cells stably expressing human $5\text{-}HT_{2B}$ receptors were used instead of transiently-transfected cells. Cells were set up and tested in the same way as above, with 5-HT serving as an agonist control (3 pM–30 µM), and with 1 nM 5-HT being used in the second addition to determine the antagonist activity of ogerin.

Phosphatidylinositol hydrolysis assay. HEK293T cells were transfected for 24 h and plated in poly-L-Lys-coated 96-well black clear bottom cell culture plates in DMEM supplemented with 10% FBS, at a density of 60,000 cells in 100 µl per well. After 5 h, cells were washed with inositol-free DMEM once and labelled with ^{3}H-inositol (1 µCi per well, PerkinElmer) in inositol-free DMEM supplemented with 5% dialysed FBS overnight. On the assay day, labelling medium was removed and cells were washed once with assay buffer (1× HBSS, 10 mM HEPES, 10 mM MES and 20 mM LiCl, pH 8.4). To measure drug concentration responses, then cells were then incubated with drug solutions at pH 8.4 for 20 min. To measure proton concentration responses, the assay buffer was pre-adjusted to desired pH values and supplemented with 20 mM LiCl. To measure the effect of ogerin or its isomer ZINC32547799 on proton concentration–response curves, pH solutions were supplemented with 20 mM LiCl and 10 µM ogerin or ZINC32547799. The premixed drug solutions were added to cells for 20 min. At the end of incubation, drug solutions were removed and 40 µl per well of 50 mM ice-cold formic acid was added. After incubation at 4 °C for 30 min, the acid extracts were transferred to polyethylene terephthalate 96-well sample plates (1450-401, Perkin Elmer) and mixed with 75 µl (200 µg) YSi RNA binding beads (RPNQ0013, Perkin Elmer). The plate was sealed and further incubated at 4 °C for 30 min before being counted on a TriLux MicroBeta counter. Results (c.p.m. per well) were analysed using Graphpad Prism.

Functional assays with A_{2A} and CXCR4 receptors. Functional assays with A_{2A} adenosine and CXCR4 chemokine receptors were carried out using a slightly different protocol from that previously described for G_s (above) and G_i receptors[56]. Specifically, HEK293T cells were transfected and plated using regular DMEM supplemented with 1% dialysed FBS. Before assays, culture medium was removed, and cells were incubated with 20 µl drug solution (prepared in drug buffer 20 mM HEPES, 1× HBSS, pH 7.4) for 15 min at room temperature. To measure

agonist activity, 5 µl of 5× luciferin solution (4 mM final concentration) for A_{2A} (G_s-coupled GPCRs) or a mixture of luciferin and isoproterenol at a final concentration of 200 nM for CXCR4 (G_i-coupled GPCRs) was added and cells were incubated for another 20 min. To measure antagonist activity, test compound was added first for 10 min before a reference agonist at a final of EC_{80} concentration for another 10 min, and then followed by addition of luciferin for A_{2A} or a mixture of luciferin and isoproterenol for CXCR4 as above. Luminescence was measured in a luminescence counter. Results were analysed in GraphPad Prism.

Radioligand binding assays. Radioligand binding assays with selected CNS targets were carried out as described[56,57] and as detailed in the PDSP protocol book available online (http://pdsp.med.unc.edu/pdspw/binding.php). In brief, receptor membrane preparations were made from either animal brain tissues, or stable cell lines, or transiently transfected HEK293T cells. Receptor expression levels and radioligand binding affinities were determined with saturation binding assays. Competition binding assays were performed with membrane aliquots and a fixed concentration of radioligand in 96-well plates in a final volume of 125 µl. Reactions were incubated in the dark and at room temperature (22 °C), and terminated by vacuum filtration onto 96-well formatted GF/B filters. Radioactivity on the filters was counted in a beta counter. Results were analysed in GraphPad Prism.

Anti-HA immunoblots. HEK293 cells were transfected with either pcDNA3 vector containing a haemagglutinin (HA) cassette within the multiple cloning site, or pcDNA3HA-GPR68 encoding human GPR68 with an N-terminal HA tag. Stable lines were generated by selection with 250 µg ml^{-1} G418, with >90% of cells expressing HA after 2 weeks as assessed by immunocytochemistry (not shown). Cells were plated into 12-well plates, grown to confluence, and media switched to Hams-F12 media, with pH adjusted to pH 8.0 or 7.4, for 1 h. Cells were then stimulated with vehicle, 50 µM ogerin, or 50 µM lorazepam for 10 min. Lysates were collected and subjected to immunoblotting, with blots probed using primary antibodies against HA (Sigma cat H3663), total vasodilator-stimulated phosphoprotein (VASP, BD Biosciences, cat 610448), p-p42/p44 (Cell Signaling, cat 5726S), and β-actin (Sigma, cat A1978), and secondary antibodies (Licor, cat 926-32213 and 926–32210) conjugated with infrared fluorophores as described previously[58].

Anti-Flag immunoblots. HEK293T cells were transiently transfected in 10-cm dishes with Flag-tagged GPR68 wild-type and mutant receptors. Untransfected HEK293T cells served as a negative control. After 48 h, cells were collected, lysed and sonicated to shear chromatin before being subjected to immunoblotting. Blots were probed with monoclonal anti-Flag M2-peroxidase antibody (Sigma, A8952). Bands were quantified and normalized to GPR68 wild-type receptor (fold) for graphing.

Data analysis and reporting. Other than *in vivo* studies (below), no statistical analysis was applied to yeast- or cell-based screening assays. Sample size (number of assays for each compound or receptor) was predetermined to be in triplicate or quadruplicate for primary screening assays at a single concentration. Some samples were repeated more than the others in the primary screening assays and the number of measurements were specified as a range in corresponding figure legends. For concentration–response assays, the sample size (number of assays for each compound at selected receptors) was also predetermined to be tested for a minimum of three assays, each in triplicate or quadruplicate. Samples or receptors were tested not randomly but in an alphabetic order or numeric order according to their coded names for easy organization and were thus blinded. For each batch of assays, a control assay with isoproterenol and proton concentration–responses were included. If potency values for either isoproterenol or proton was >0.5 log unit away from established averages, assays with the batch of transfected cells were excluded. For structure–activity relationship (SAR) studies, only the assays in which all related compounds were tested side by side were included. None of the functional assays were blinded to investigators.

Generation of GPR68-knockout mice. To generate GPR68-knockout mice, a probe specific for the human *GPR68* transcript was generated by PCR amplification of a 450-base-pair (bp) segment of the coding sequence of the final exon of *GPR68* using total placental RNA. The probe was used to identify a clone from a 129 mouse genomic lambda library. The genomic insert was subcloned and a restriction map generated using a panel of enzymes. The targeting construct for the *GPR68* locus consists of a PGK-1 promoter driven neomycin resistance cassette flanked by two arms of homology with the mouse *GPR68* locus. The longer arm of homology was generated using a 7,266-bp PstI fragment extending from the last intron to the beginning of the last exon. This exon contains the entire coding sequence of the *GPR68* gene. The 1,335-bp shorter arm was generated by PCR amplification and extends from the downstream end of the long arm into the 3′ untranslated region of the gene. Homologous recombination of the targeting construct with the *GPR68* locus inserts the neomycin resistance cassette into codon 78 of the gene, thereby disrupting expression. Correctly targeted cell lines were identified by Southern blot analysis using a probe consisting of a 1,496-bp PstI fragment immediately upstream of the long arm. This probe recognizes a 14,290-bp EcoRV fragment in the endogenous locus and a 7,855-bp fragment in the targeted locus. Genotyping was carried out by PCR with three primers. The common (5′-GCAG AGGAAGCCCACGCTGATGTA-3′) and endogenous (5′-TAAACGGTAGCTGT GATTATTCAA-3′) primers generate a 516-bp PCR product from the endogenous locus, while the common and targeted (5′-AAATGCCTGCTCTTTACTGAAGG-3′) primers generate a 465-bp product from the targeted locus. The chimaeras were bred to C57BL/6J mice and pups carrying the mutant allele identified. After ten successive crosses of heterozygous animals to C57BL/6J mice, heterozygous mice were intercrossed and a congenic *Gpr68*$^{-/-}$ and C57BL/6J breeding colony established. The GPR68-knockout mice were profiled in several behavioural tests as described below in detail and results are summarized in Extended Data Fig. 11 and Supplementary Tables 11 and 12.

***In vivo* behavioural profiles of GPR68-knockout mice.** Mice were maintained and handled according to the Guide for the Care and Use of Laboratory Animals approved by the Institutional Animal Care and Use Committee of the University of North Carolina at Chapel Hill. The goal of this study was to determine whether targeted deletion of GPR68 alters behavioural function in mice.

Timeline for behavioural tests. The following tests were performed with mice at the ages shown in parentheses. Elevated plus maze test for anxiety-like behaviour (6–7 weeks); activity in an open field, accelerating rotarod (2 tests, 48 h apart) (7–8 weeks); three-chamber social approach test, activity in an open field (re-test) (8–9 weeks); marble-burying assay (9–10 weeks); acoustic startle test, buried food test for olfactory ability (10–11 weeks); visual cue test in the Morris water maze (11–12 weeks); hidden platform test for spatial learning (12–14 weeks); reversal learning in the Morris water maze (14–16 weeks); second acoustic startle test, hotplate test for thermal sensitivity (16–17 weeks).

Summary of results. Mice with deletion of GPR68 had normal performance in most of the behavioural tests. No effects of genotype were observed for body weights, activity and anxiety-like behaviour in an elevated plus maze or an open field, motor coordination, sociability, prepulse inhibition of acoustic startle responses or acquisition in the water maze. However, both male and female GPR68-knockout mice had small, significant decreases in acoustic startle responses, suggesting a reduced responsivity to environmental stimuli. Male GPR68-knockout mice also showed significant decreases in marble burying, a test for anxiety-like phenotypes. Overall, the findings indicate that GPR68 might have a role in specific domains of behaviour.

Elevated plus maze. This test is used to assess anxiety-like behaviour in rodents. The procedure is based on a natural tendency of mice to actively explore a new environment, versus a fear of being in an open area. In the present study, mice were given one 5-min trial on the plus maze, which had two walled arms (the closed arms, 20 cm in height) and two open arms. The maze was elevated 50 cm from the floor, and the arms were 30 cm long. Animals were placed on the centre section (8 × 8 cm), and allowed to freely explore the maze. Measures were taken of time on, and number of entries into, the open and closed arms. All of the experimental groups showed a strong preference for the closed arms, in comparison to the open arms, of the elevated plus maze. As shown in Supplementary Table 11, there were no significant differences between the wild-type and GPR68-knockout mice for percentage time or percentage entries on the open arms, or for total entries during the task.

Activity in an open field. Exploratory activity in a novel environment was assessed in an open field chamber (41 × 41 × 30 cm) crossed by a grid of photobeams (VersaMax system, AccuScan Instruments). Counts were taken of the number of photobeams broken during the trial in 5-min intervals, with separate measures for ambulation (total distance travelled) and rearing movements. Time spent in the centre region of the open field was measured as an index of anxiety-like behaviour. Unfortunately, an equipment malfunction led to the loss of data for 8 mice during the first activity test, conducted when mice were 7–8 weeks in age. Therefore, a second activity test was given, when mice were 8–9 weeks in age. As depicted in Extended Data Fig. 11a, b, there were no significant differences between the wild-type and GPR68-knockout mice for distance travelled, or for rearing or centre time (data not shown), during the second activity test. A significant sex × time interaction was found for the distance measure ($F_{(11,385)} = 2.68$, $P = 0.0025$), reflecting higher levels of activity in the female groups at the beginning of the session.

Accelerating rotarod test. Subjects were tested for motor coordination and learning on an accelerating rotarod (Ugo Basile). For the first test session, animals were given three trials, with 45 s between each trial. Two additional trials were given 48 h later. Revolutions per minute (rpm) was set at an initial value of 3, with a progressive increase to a maximum of 30 rpm across five minutes (the maximum trial length). Measures were taken for latency to fall from the top of the rotating barrel. As shown in Extended Data Fig. 11c, d, deletion of GPR68 did not lead to deficits in motor coordination on the rotarod. In fact, during the first three

acquisition trials, there was a non-significant trend for enhanced performance in the male knockout group (repeated-measures ANOVA, genotype × sex interaction, $F_{(1,35)} = 3.58$, $P = 0.0668$).

Marble-burying assay. This procedure is used to evaluate anxiety-like behaviour and repetitive responses. Mice were tested in a Plexiglas cage located in a sound-attenuating chamber with ceiling light and fan. The cage contained 5 cm of corncob bedding, with 20 black glass marbles (14 mm diameter) arranged in an equidistant 5 × 4 grid on top of the bedding. Animals were given access to the marbles for 30 min. Measures were taken of the number of buried marbles (two-thirds of the marble covered by the bedding). A two-way ANOVA indicated a significant genotype × sex interaction ($F_{(1,35)} = 7.37$, $P = 0.0102$) (Supplementary Table 11). Post-hoc comparisons revealed that the male GPR68-knockout mice buried significantly fewer marbles than both male wild-type mice and female knockout mice in this task.

Buried food test for olfactory function. Several days before the olfactory test, an unfamiliar food (Froot Loops, Kellogg Co.) was placed overnight in the home cages of the mice. Observations of consumption were taken to ensure that the novel food was palatable. Sixteen to twenty hours before the test, all food was removed from the home cage. On the day of the test, each mouse was placed in a large, clean tub cage (46 × 23.5 × 20 cm (width, length, height)), containing paper chip bedding (3-cm deep), and allowed to explore for 5 min. The animal was removed from the cage, and one Froot Loop was buried in the cage bedding. The animal was then returned to the cage and given fifteen minutes to locate the buried food. Measures were taken of latency to find the food reward. As shown in Supplementary Table 11, there were no significant differences between the groups in latency to find the buried food.

Hotplate test for thermal sensitivity. Individual mice were placed in a tall plastic cylinder located on a hotplate, with a surface heated to 55 °C (IITC Life Science). Reactions to the heated surface, including hindpaw lick, vocalization or jumping, led to immediate removal from the hotplate. Measures were taken of latency to respond. The maximum test length was 30 s, to avoid paw damage. A two-way ANOVA indicated a significant main effect of sex ($F_{(1,1)} = 8.83$, $P = 0.0053$), and genotype × sex interaction ($F_{(1,35)} = 4.3$, $P = 0.0455$) (Supplementary Table 11). Post-hoc comparisons revealed that the male GPR68-knockout mice had significantly lower latencies to respond than female knockout mice.

Acoustic startle method. The acoustic startle test can be used to assess auditory function and sensorimotor gating. The test is based on the measurement of the reflexive whole-body flinch, or startle response, that follows exposure to a sudden noise. Mice can be evaluated for levels of startle magnitude and prepulse inhibition, which occurs when a weak prestimulus leads to a reduced startle in response to a subsequent louder noise. For this study, animals were tested with a San Diego Instruments SR-Lab system. In brief, mice were placed in a small Plexiglas cylinder within a larger, sound-attenuating chamber. The cylinder was seated upon a piezoelectric transducer, which allowed vibrations to be quantified and displayed on a computer. The chamber included a house light, fan, and a loudspeaker for the acoustic stimuli. Background sound levels (70 dB) and calibration of the acoustic stimuli were confirmed with a digital sound level meter (San Diego Instruments). Each session consisted of 42 trials, which began with a 5-min habituation period. There were seven different types of trials: the no-stimulus trials, trials with the acoustic startle stimulus (40 ms; 120 dB) alone, and trials in which a prepulse stimulus (20 ms; 74, 78, 82, 86 or 90 dB) occurred 100 ms before the onset of the startle stimulus. Measures were taken of the startle amplitude for each trial across a 65-ms sampling window, and an overall analysis was performed for each subject's data for levels of prepulse inhibition at each prepulse sound level (calculated as 100−(response amplitude for prepulse stimulus and startle stimulus together/response amplitude for startle stimulus alone) ×100).

Results from acoustic startle test. The GPR68-knockout mice had decreased startle responses after presentation of acoustic stimuli, in comparison to the wild-type mice (Extended Data Fig. 11e, f). A repeated-measures ANOVA, conducted on startle response amplitudes, indicated significant main effects of genotype ($F(1,35) = 7.22$, $P = 0.011$) and sex ($F(1,35) = 16.61$, $P = 0.0003$), and a genotype × decibel level interaction ($F(6,210) = 5.77$, $P < 0.0001$). Separate comparisons confirmed that both male and female knockout mice showed significant reductions in startle responses (genotype × decibel level interaction, males, $F(6,84) = 2.57$, $P = 0.0245$; and females, $F(6,126) = 3.48$, $P = 0.0032$). The decreased startle responses and overt sex differences were not associated with changes in prepulse inhibition (Extended Data Fig. 11g, h). The significant main effects of genotype on startle were no longer evident during a second acoustic startle test, conducted when mice were 16–17 weeks in age.

Morris water maze, visible platform test. The Morris water maze task was used to assess spatial learning and visual function in the mice. The water maze consisted of a large circular pool (diameter = 122 cm) partially filled with water (45 cm deep, 24–26 °C), located in a room with numerous visual cues. Mice were first tested using a visible platform. In this case, each animal was given four trials per day, across 2 days, to swim to an escape platform cued by a patterned cylinder extending above the surface of the water. For each trial, the mouse was placed in the pool at one of four possible locations (randomly ordered), and then given 60 s to find the visible platform. If the mouse found the platform, the trial ended, and the animal was allowed to remain 10 s on the platform before the next trial began. If the platform was not found, the mouse was placed on the platform for 10 s, and then given the next trial. Measures were taken of latency to find the platform via an automated tracking system (Noldus Ethovision). As shown in Supplementary Table 12, all groups of mice demonstrated a high degree of proficiency in the visual cue task.

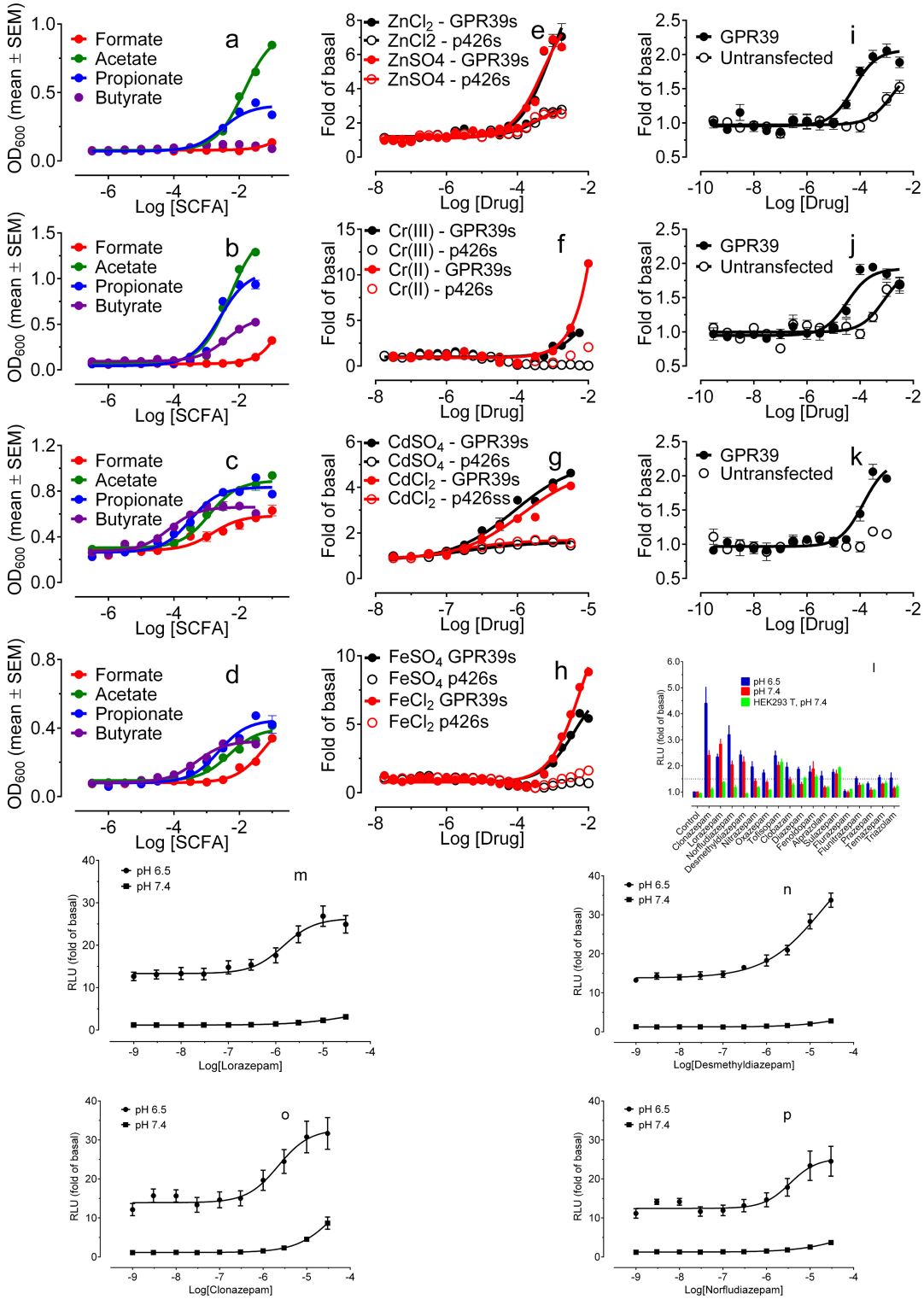
Acquisition and reversal learning in the hidden platform test (Extended Data Fig. 11i–l). Three days after the visual cue task, mice were tested for their ability to find a submerged, hidden escape platform (diameter = 12 cm). As in the procedure for visual cue learning, each animal was given four trials per day, with 1-min per trial, to swim to the hidden platform. The criterion for learning was an average latency of 15 s or less to locate the platform on 1 day. Mice were tested until the criterion was reached, with a maximum of 9 days of testing. When criterion was reached, mice were given a 1-min probe trial in the pool with the platform removed. In this case, selective quadrant search was evaluated by measuring number of crosses over the location where the platform (the target) had been placed during training, and the corresponding areas in the other three quadrants. After the acquisition phase, mice were tested for reversal learning, using the same procedure as described above. In this phase, the hidden platform was located in a different quadrant in the pool, diagonal to its previous location. As before, measures were taken of latency to find the platform. On the day that the criterion for learning was met, the platform was removed from the pool, and the group was given a probe trial to evaluate reversal learning.

For the above behavioural profiling studies, subjects were 21 wild-type mice (9 males and 12 females) and 18 GPR68-knockout mice (7 males and 11 females), on a C57BL/6 background. Sample sizes were not statistically predetermined. Testing began when animals were 6–7 weeks of age. For each procedure, measures were taken by an observer blinded to mouse genotype (wild type or knockout) and no animals were excluded from analysis. Data were analysed using one-way or repeated-measures ANOVA. Fisher's protected least-significant difference tests were used for comparing group means only when a significant F value was determined. Within-group comparisons were conducted to determine side preference in the social behaviour tests. For all comparisons, significance was pre-set at $P < 0.05$.

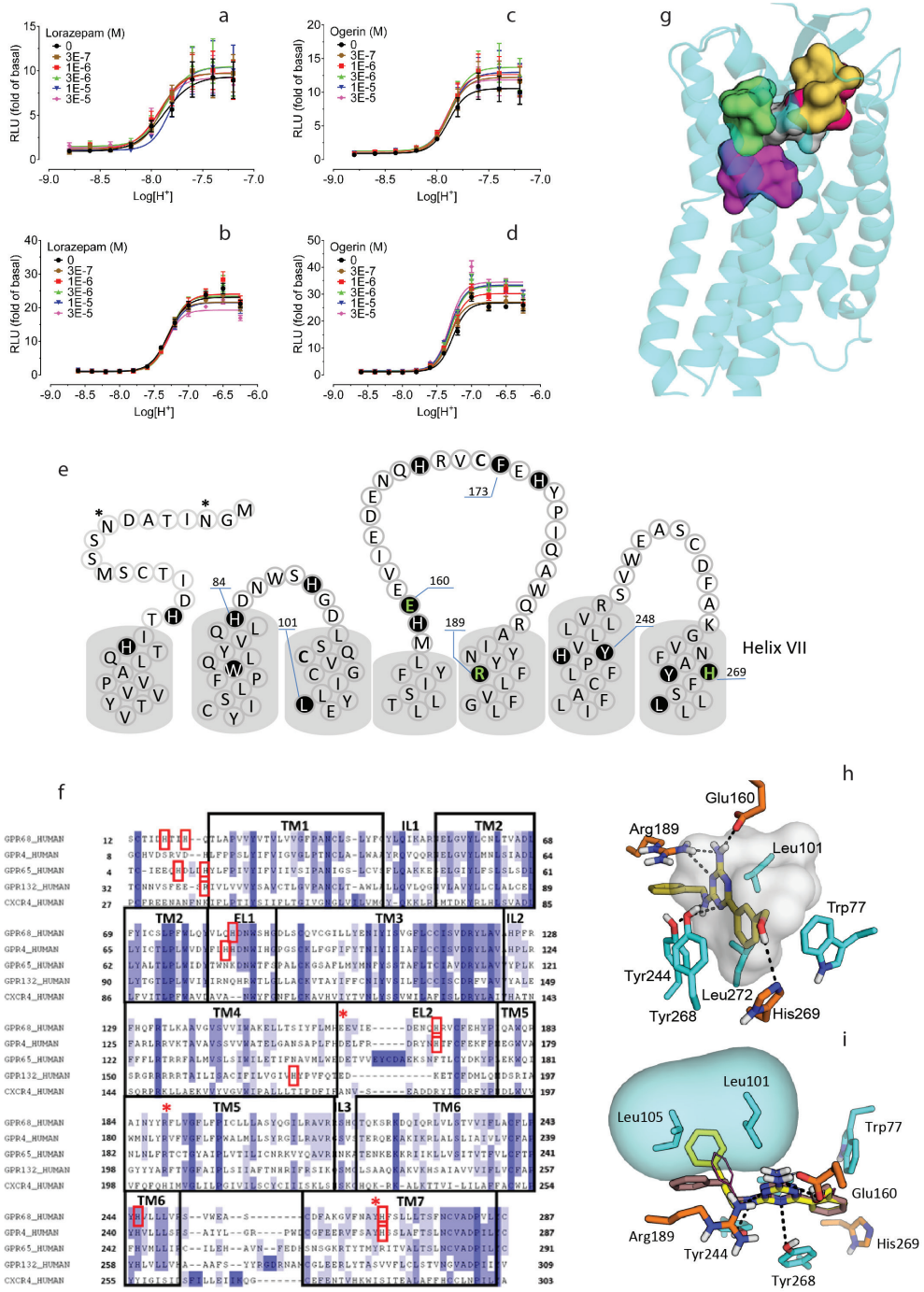
Effect of ogerin and its analogue ZINC32547799 on learning and memory. Contextual and cue-dependent learning and memory were evaluated using a Near-Infrared Video Fear Conditioning system (MED Associates). Test chambers (29 × 25 × 25 cm) had transparent walls and metal rod floors, and were enclosed in sound-attenuating boxes. The conditioned fear procedure had three phases: training, a test for contextual learning, and a test for cue-dependent learning. Before each phase, mice were moved to a holding room adjacent to the test room and acclimated for at least 30 min. In the 8-min training phase, mice receive three pairings of a 30-s, 90-dB, 5-kHz tone (the conditioned stimulus) and a 2-s, 0.6-mA foot shock (the unconditioned stimulus), in which the shock was presented during the last 2 s of the tone. Context-dependent learning was evaluated 24 h after the training phase. Mice were placed back into the original test chamber, and levels of freezing (immobility) were determined across a 5-min session, without the presence of the conditioned or unconditioned stimulus. Forty-eight hours after the training phase, mice were evaluated for associative learning to the auditory cue (the conditioned stimulus) in a final 6-min session. The conditioning chambers were modified using a Plexiglas insert to change the wall and floor surface, and a novel odour (vanilla flavouring) was added to the sound-attenuating box. Baseline behaviour was scored for 2 min, and then three 30-s conditioned stimulus tones were presented across a 4-min period. Levels of freezing were automatically measured by the image tracking software (Med Associates). Freezing was defined as no movement (below the movement threshold) for 0.5 s. To evaluate the effect of drug, strain-matched group of animals were given ogerin (10 mg kg^{-1} in 10% Tween 80 or saline) 30 min before the training.

For the learning and memory studies, sample sizes (number of animals) were not predetermined by a statistical method, and minimum of six male animals (age of 6–8 weeks) were used in each group (exact number of animals was specified in figure legends). Animals were assigned to groups randomly and experiments were not blinded to investigators. No animals were excluded from analysis. Statistical analyses were performed after first assessing the normality of distributions of data sets. Comparisons between groups were made using unpaired t-tests. Welch's corrections were used when variances between groups were unequal. Comparisons between groups during conditioning, contextual and cued memory tests were assessed using two-way ANOVA with $P < 0.05$ being considered significant.

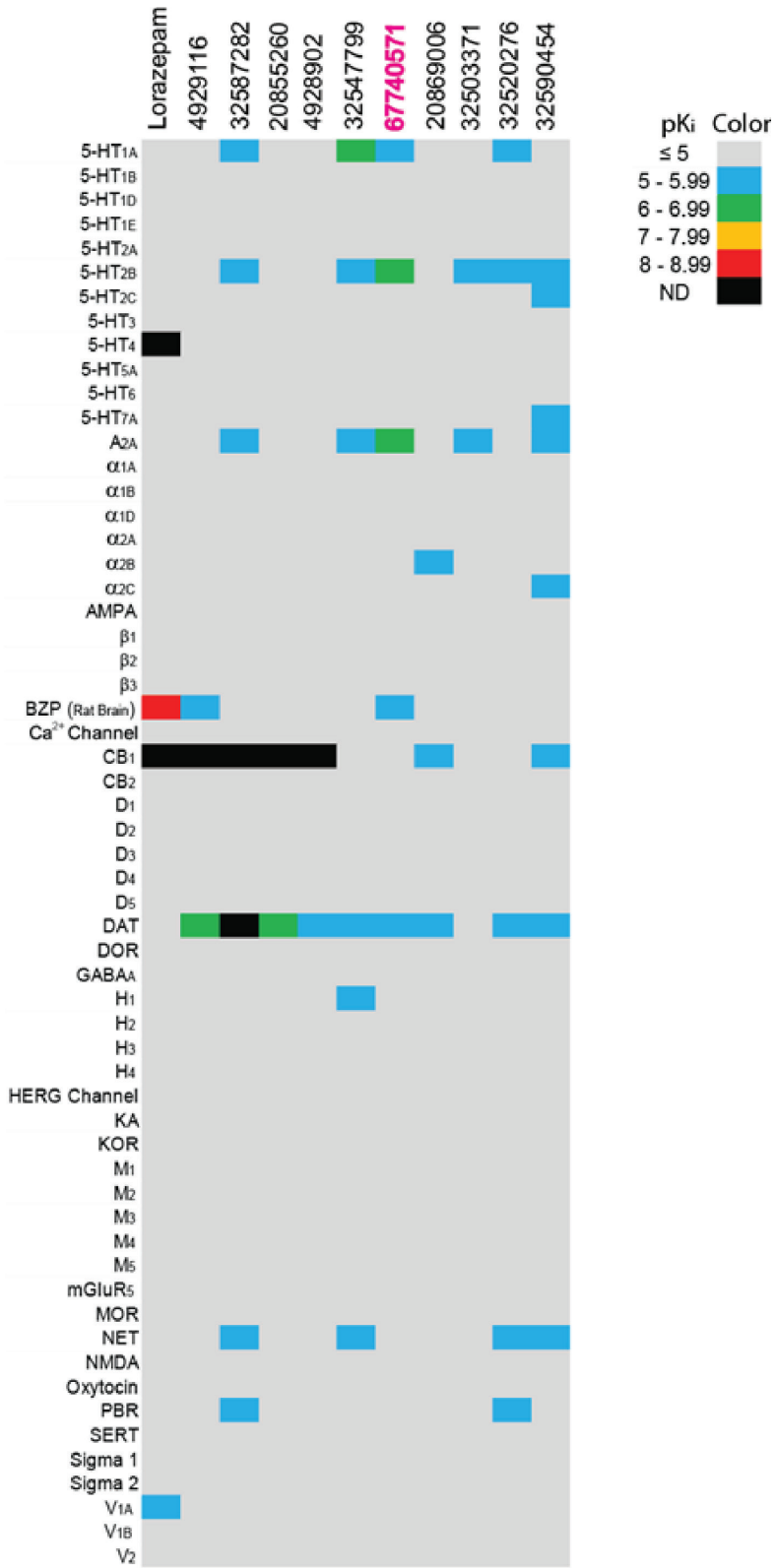
45. Eswar, N. *et al.* Comparative protein structure modeling using MODELLER. *Curr. Protoc. Protein Sci.* **Chapter 2,** Unit 2.9 (2001).
46. Yang, Q. & Sharp, K. A. Building alternate protein structures using the elastic network model. *Proteins* **74,** 682–700 (2009).
47. Jacobson, M. P., Friesner, R. A., Xiang, Z. & Honig, B. On the role of the crystal environment in determining protein side-chain conformations. *J. Mol. Biol.* **320,** 597–608 (2002).
48. Li, J., Zhu, T., Cramer, C. J. & Truhlar, D. G. New class IV charge model for extracting accurate partial charges from wave functions. *J. Phys. Chem. A* **102,** 1820–1831 (1998).
49. Chambers, C. C., Hawkins, G. D., Cramer, C. J. & Truhlar, D. G. Model for aqueous solvation based on class IV atomic charges and first solvation shell effects. *J. Phys. Chem.* **100,** 16385–16398 (1996).
50. Hert, J., Keiser, M. J., Irwin, J. J., Oprea, T. I. & Shoichet, B. K. Quantifying the relationships among drug classes. *J. Chem. Inf. Model.* **48,** 755–765 (2008).
51. Gaulton, A. *et al.* ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* **40,** D1100–D1107 (2012).
52. Mumberg, D., Muller, R. & Funk, M. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* **156,** 119–122 (1995).
53. Erlenbach, I. *et al.* Functional expression of M_1, M_3 and M_5 muscarinic acetylcholine receptors in yeast. *J. Neurochem.* **77,** 1327–1337 (2001).
54. Armbruster, B. N., Li, X., Pausch, M. H., Herlitze, S. & Roth, B. L. Evolving the lock to fit the key to create a family of G protein-coupled receptors potently activated by an inert ligand. *Proc. Natl Acad. Sci. USA* **104,** 5163–5168 (2007).
55. Christopoulos, A. & Kenakin, T. G protein-coupled receptor allosterism and complexing. *Pharmacol. Rev.* **54,** 323–374 (2002).
56. Besnard, J. *et al.* Automated design of ligands to polypharmacological profiles. *Nature* **492,** 215–220 (2012).
57. Keiser, M. J. *et al.* Predicting new molecular targets for known drugs. *Nature* **462,** 175–181 (2009).
58. Horvat, S. J. *et al.* A-kinase anchoring proteins regulate compartmentalized cAMP signaling in airway smooth muscle. *FASEB J.* **26,** 3670–3679 (2012).
59. Huang, X.-P., Mangano, T., Hufeisen, S., Setola, V. & Roth, B. L. Identification of human *ether-à-go-go* related gene modulators by three screening platforms in an academic drug-discovery setting. *Assay Drug Dev. Technol.* **8,** 727–742 (2010).



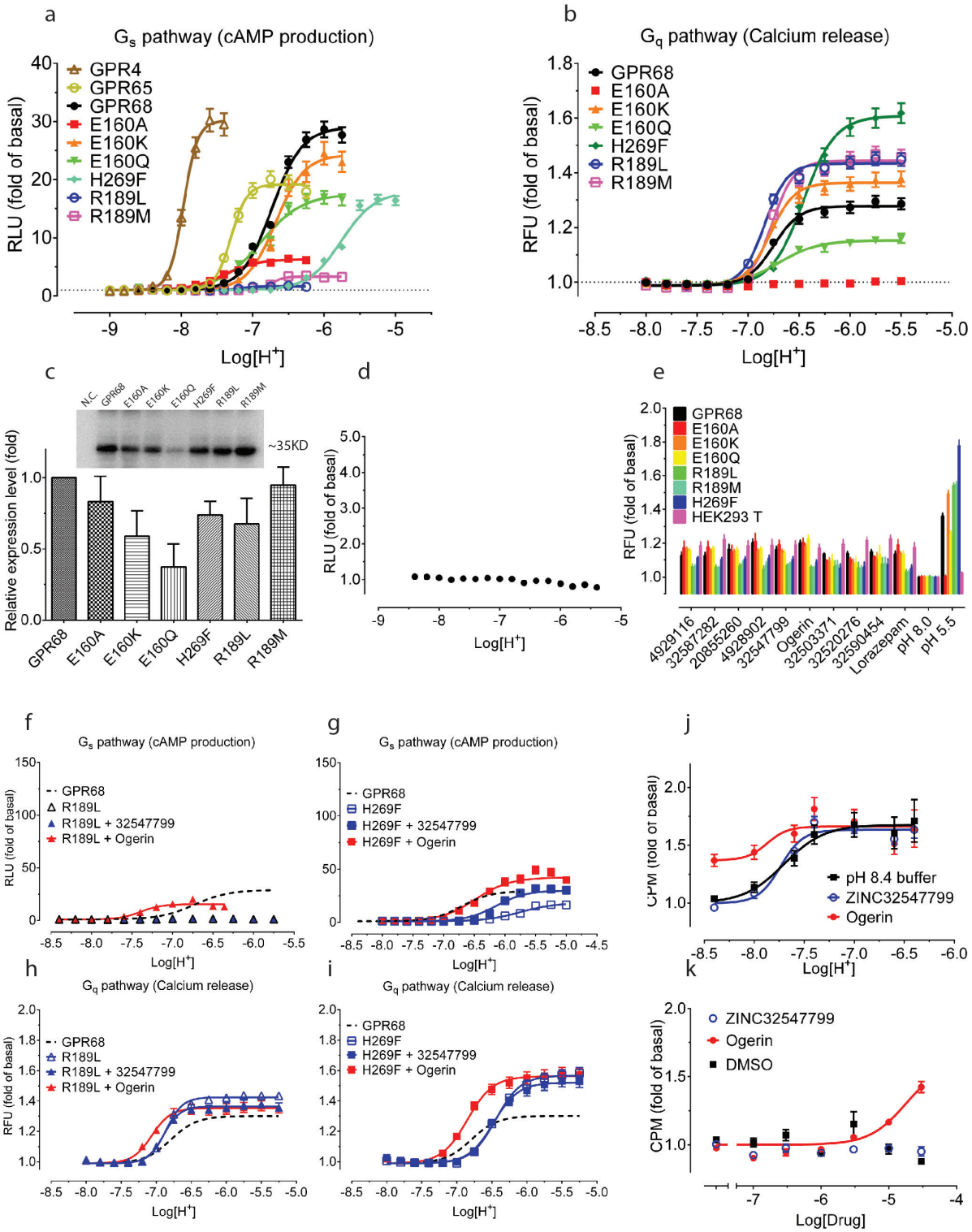
Extended Data Figure 1 | Validation and confirmation of GPCR activation assays. a–o, Yeast (**a–k**) and HEK293T cell (**l–o**) GPCR activation assays. **a–d**, Concentration-dependent growth of GPR43-expressing G_s yeast (**a**), GPR43-expressing G_q yeast (**b**), GPR41-expressing G_s yeast (**c**), and GPR41-expressing G_q yeast (**d**) in response to various short-chain fatty acids (SCFAs). **e–h**, Concentration-dependent growth of GPR39-expressing G_s yeast ($GPR39_s$) in response to zinc ions (**e**), chromium ions (**f**), cadmium ions (**g**) and iron ions (**h**). **i–k**, Concentration-dependent cAMP responses of GPR39-expressing HEK293T cells to $ZnCl_2$ (**i**), $ZnSO_4$ (**j**) or $CdSO_4$ (**k**) as measured by luciferase cAMP reporter assay. **l**, *N*-unsubstituted benzodiazepines (lorazepam, clonazepam, desmethyldiazepam and norfludiazepam; 10 μM) stimulated cAMP production in a GPR68- and pH-dependent manner. Data are mean ± s.e.m. ($n = 3$–66 measurements). **m–p**, Concentration–response curves of *N*-unsubstituted benzodiazepines lorazepam (**m**), desmethyldiazepam (**n**), clonazepam (**o**) and norfludiazepam (**p**) at pH 6.50 or 7.40 in GPR68-transfected HEK293T cells (structures in Supplementary Table 1). Normalized results represent mean ± s.e.m. ($n = 3$) and curves were analysed in GraphPad Prism using the built-in 4 parameter logistic function.



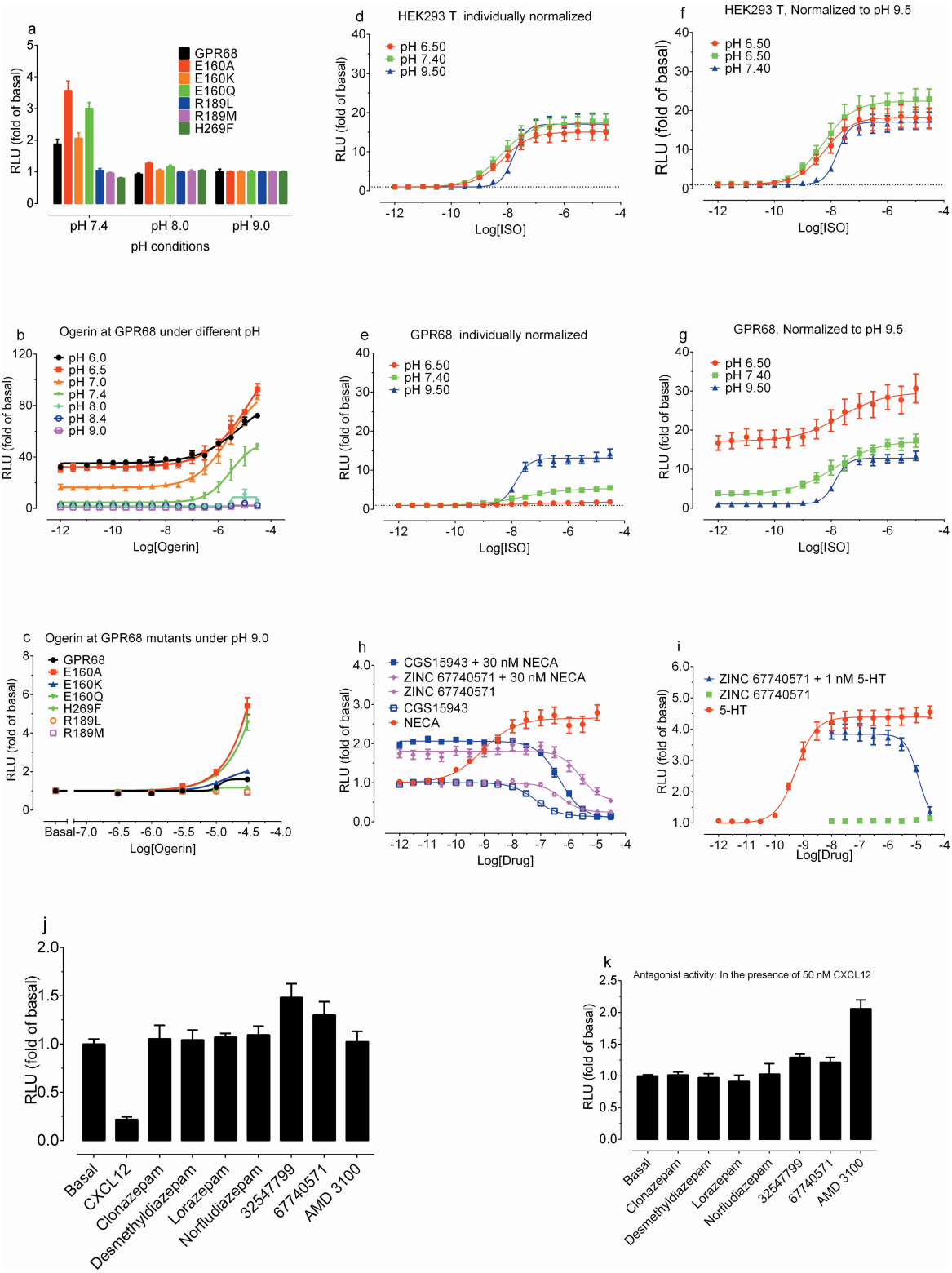
Extended Data Figure 2 | Lorazepam and ogerin have minimal GPR4 or GPR65 activity. a–d, Effect of lorazepam (**a**, **b**) or ogerin (**c**, **d**) on GPR4 (**a**, **c**) or GPR65 (**b**, **d**); data represent normalized mean ± s.e.m. ($n = 3$). Sequencing alignment proton-sensing receptor and docking poses for ogerin and its analogues. **e**, GPR68 snake plot showing extracellular loops and transmembrane domains (upper portion); important residues are highlighted. Glu160, Arg189 and His269 were mutated in this study. **f**, Sequence alignment of GPR4, GPR65 and GPR68 to CXCR4 (PDB code 3ODU) (PROMALS-3D) was manually refined to reduce gaps and to position conserved residues. TM, transmembrane regions; IL, intracellular loop; EL, extracellular loop. Conserved residues highlighted in blue by degree of conservation while red boxes indicate residues important for receptor function. Red stars indicate residues mutated in this study. **g**, Sampling different regions for lorazepam binding modes in GPR68. Yellow and grey surfaces contour the binding site of 1T1t and CVX15 in CXCR4 crystal structures (PDB codes 3ODU and 3OE0, respectively), while green and red surfaces sample the entire binding pocket. The magenta surface represents the canonical orthosteric biogenic amine site. **h**, ZINC32547799 in its predicted orientation and interactions with GPR68. **i**, Optimization of ogerin (magenta, thin lines) to **C2** (brown, structure in Fig. 3a) by insertion of a single methylene is predicted to improve packing in the aryl pocket of the ogerin site. Adding a second methylene, thus creating a propyl linker in **C3** (yellow, structure in Fig. 3a), is predicted to disrupt the packing and thus to reduce the allosteric effect.



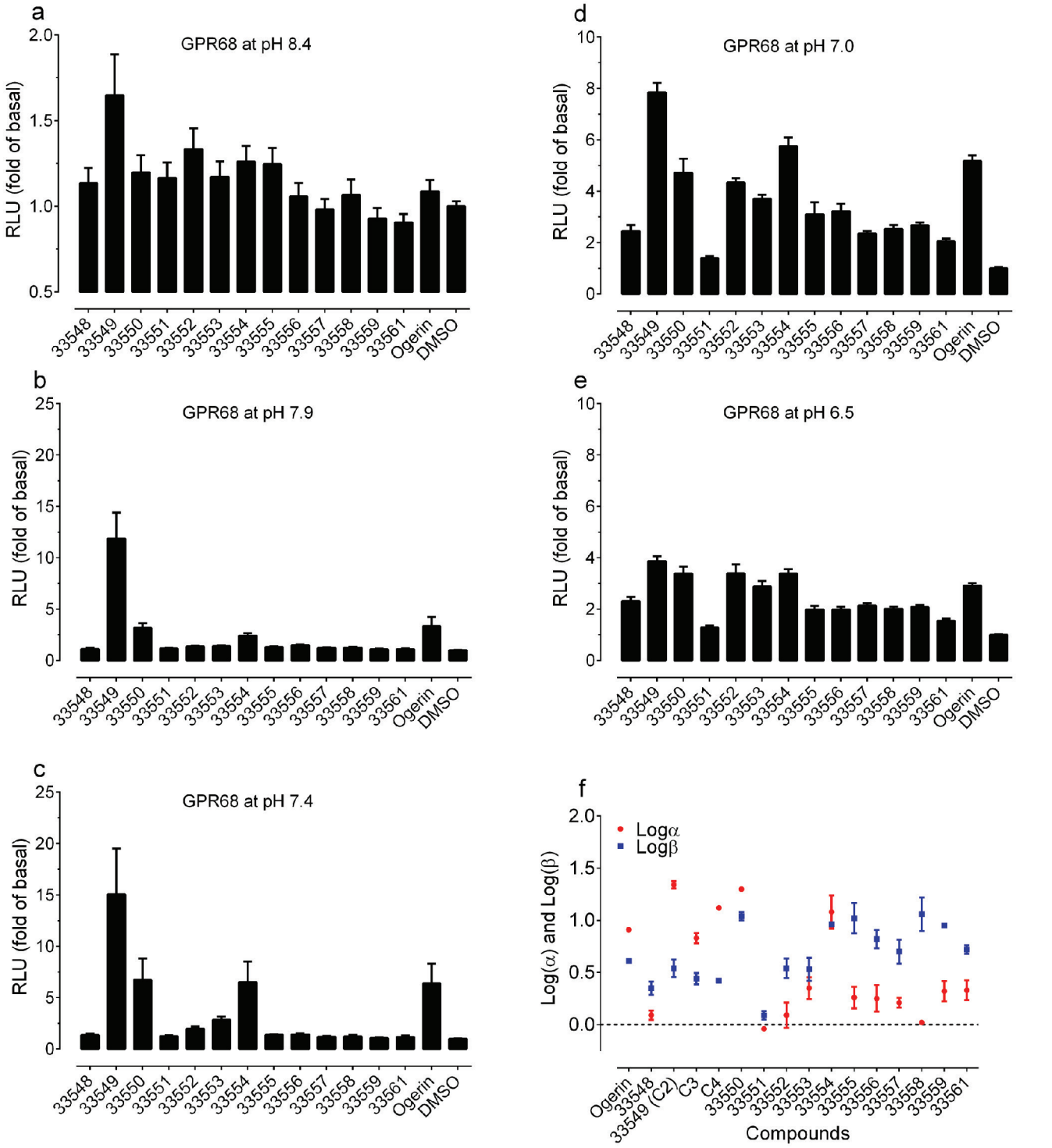
Extended Data Figure 3 | Heat map of off-target activities of lead compounds at potential CNS drug targets. Radioligand binding assays were carried out by the National Institute of Mental Health Psychoactive Drug Screening Program (NIMH PDSP) as described previously[56,57] (online protocols available at http://pdsp.med.unc.edu/pdspw/binding.php). Values represent mean binding affinities (pK_i, $n = 2–4$). Affinities lower than a pK_i of 5, or less than 50% inhibition at 10 μM, are shown as a minimum of 5 on the pK_i scale. The hERG inhibition activity was tested in a hERG functional assay as previously published[59]. AMPA, aminomethylphosphonic acid receptor; BZP, benzodiazepine receptor; DAT, dopamine transporter; DOR, delta (δ) opioid receptor; KA, kainate acid receptor; KOR, kappa (κ) opioid receptor; MOR, mu (μ) opioid receptor; NAT, noradrenaline transporter; NMDA, *N*-methyl-D-aspartate receptor; ND, not determined; PBR, peripheral benzodiazepine binding site; SERT, serotonin transporter; hERG, human ether-a-go-go-related gene (potassium channel Kv11.1).



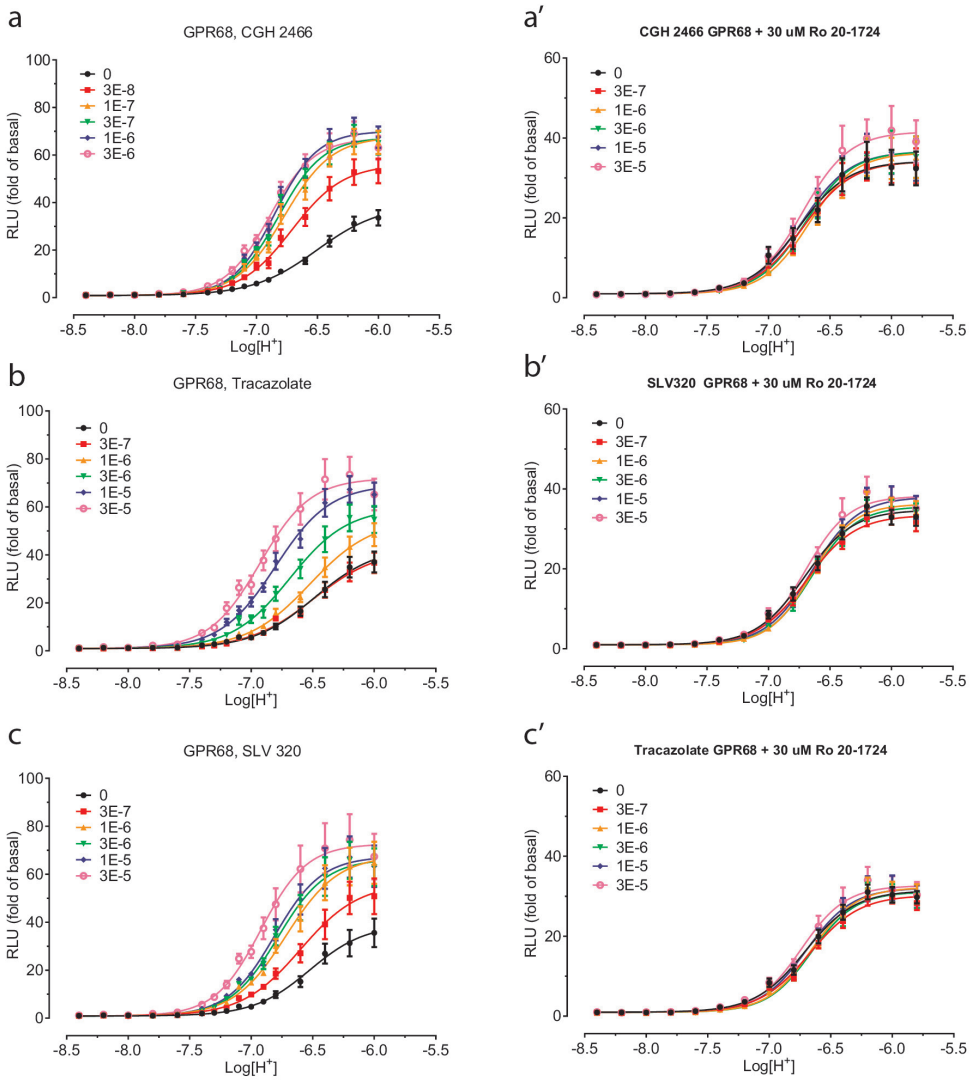
Extended Data Figure 4 | Confirmation of modelling results via mutagenesis. a, **b**, Protons showed agonist activity at GPR68 wild-type and mutant receptors in cAMP production (**a**) and calcium release (**b**); parameters are in Supplementary Table 4. **c**, Relative GPR68 wild-type and mutant receptor expression levels determined by anti-Flag immunoblotting ($n = 3$). **d**, Proton-mediated cAMP production in untransfected cells ($n = 16$). **e**, Calcium release by lorazepam and selected ZINC compounds (10 μM at pH 8.0, $n = 6$–22 measurements). **f**–**j**, Effect of ogerin and ZINC32547799 (10 μM) on proton-mediated cAMP production (**f** and **g**, $n = 4$), calcium release (**h** and **i**, $n = 3$), and phosphatidylinositol hydrolysis (**j**, $n = 3$) at GPR68 wild-type or mutant-transfected HEK293T cells. **k**, Effect of ogerin and ZINC32547799 on phosphatidylinositol hydrolysis at pH 8.4 at GPR68-transfected GPR68 HEK293T cells ($n = 3$). Normalized results represent mean ± s.e.m. and curves were analysed using a four-parameter logistic function.

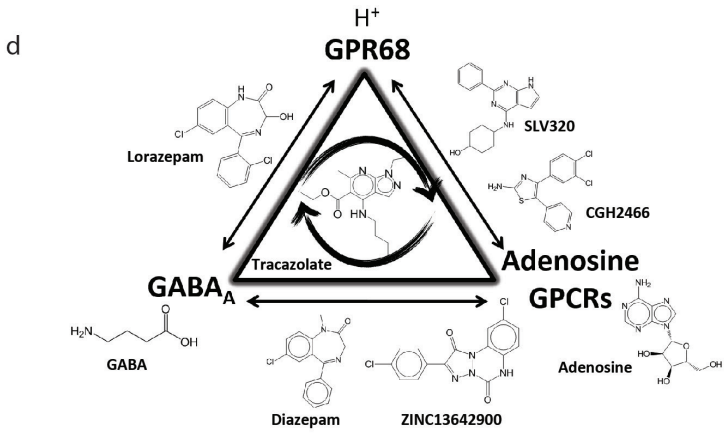


Extended Data Figure 5 | Control experiments for signalling and pharmacology. a, Basal cAMP production of GPR68 wild-type and mutant receptors (mean ± s.e.m., $n = 24$–46 measurements). **b**, pH-dependent activity of ogerin at GPR68 wild type (mean ± s.e.m., $n = 3$). **c**, Ogerin concentration–responses at GPR68 wild-type and mutant receptors at pH 9.0 (**c**, mean ± s.e.m., $n = 3$), under which cAMP reporter assay was *n*ot affected (**d**–**f**). **d**–**g**, Proton modulated isoproterenol-mediated G_s-activation via β_2-adrenergic receptors in untransfected (**d**, **f**) and GPR68-transfected (**e**, **g**) cells. Normalized results (basal at pH 9.5 for **d** and **e**; or corresponding buffer control for **f** and **g**) represent mean ± s.e.m. ($n = 6$). **h**, **i**, Inverse agonist and antagonist activity (K_i of 220 nM) of ogerin at A_{2A} (cAMP production, **h**) and weak antagonist activity (K_i of 736 nM) at 5-HT_{2B} receptors (calcium mobilization, **i**). 5′-*N*-ethylcarboxamidoadenosine (NECA) and 2-chloro-N^6-cyclopentyladenosine (CCPA) served as agonist controls, while CGS15943 is an inverse agonist control for A_{2A} receptors. Normalized results represent mean ± s.e.m. ($n = 3$). Curves were analysed in GraphPad Prism with the built-in four-parameter logistic function. **j**, **k**, Lead compounds (10 μM) showed no agonist (**j**) or antagonist (**k**) activity at CXCR4 receptors (cAMP production) with CXCL12 as an agonist control (1 or 3 μM) or AMD 3100 (10 μM) as an antagonist control. Results represent mean ± s.d. ($n = 2$).

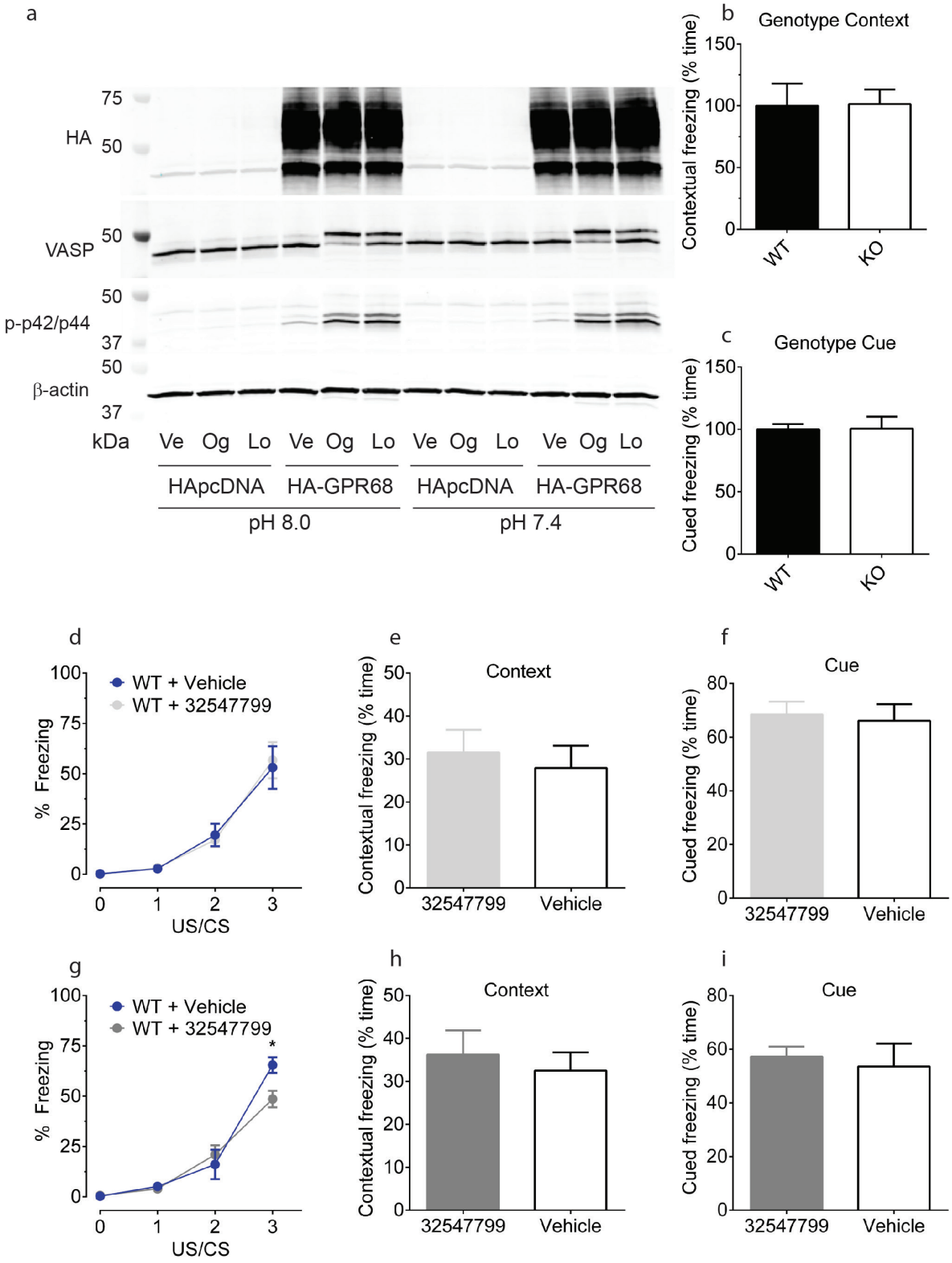


Extended Data Figure 6 | Primary screening and comparison of allosteric parameters of 13 ogerin analogues at GPR68. The 13 ogerin analogues (structures in Supplementary Table 9) identified from docking a virtual library of more than 600 ogerin derivatives were synthesized (Supplementary Information). **a**–**e**, Production of cAMP was measured in transiently transfected HEK293T cells at 10 μM and five different pH conditions, pH 8.4 (**a**); pH 7.9 (**b**); pH 7.4 (**c**); pH 7.0 (**d**); and pH 6.5 (**e**), to reveal any pH-dependent potentiation activity. Normalized results represent mean ± s.e.m. ($n = 8$–16 measurements). **f**, Graphic comparison of the allosteric parameters logα and logβ. Proton concentration–responses were carried out in the absence and presence of increasing concentrations of ogerin and its analogues, results were analysed using a standard allosteric operational model to obtain allosteric parameters. Values represent mean ± s.e.m. ($n \geq 3$; see details in Supplementary Table 8).

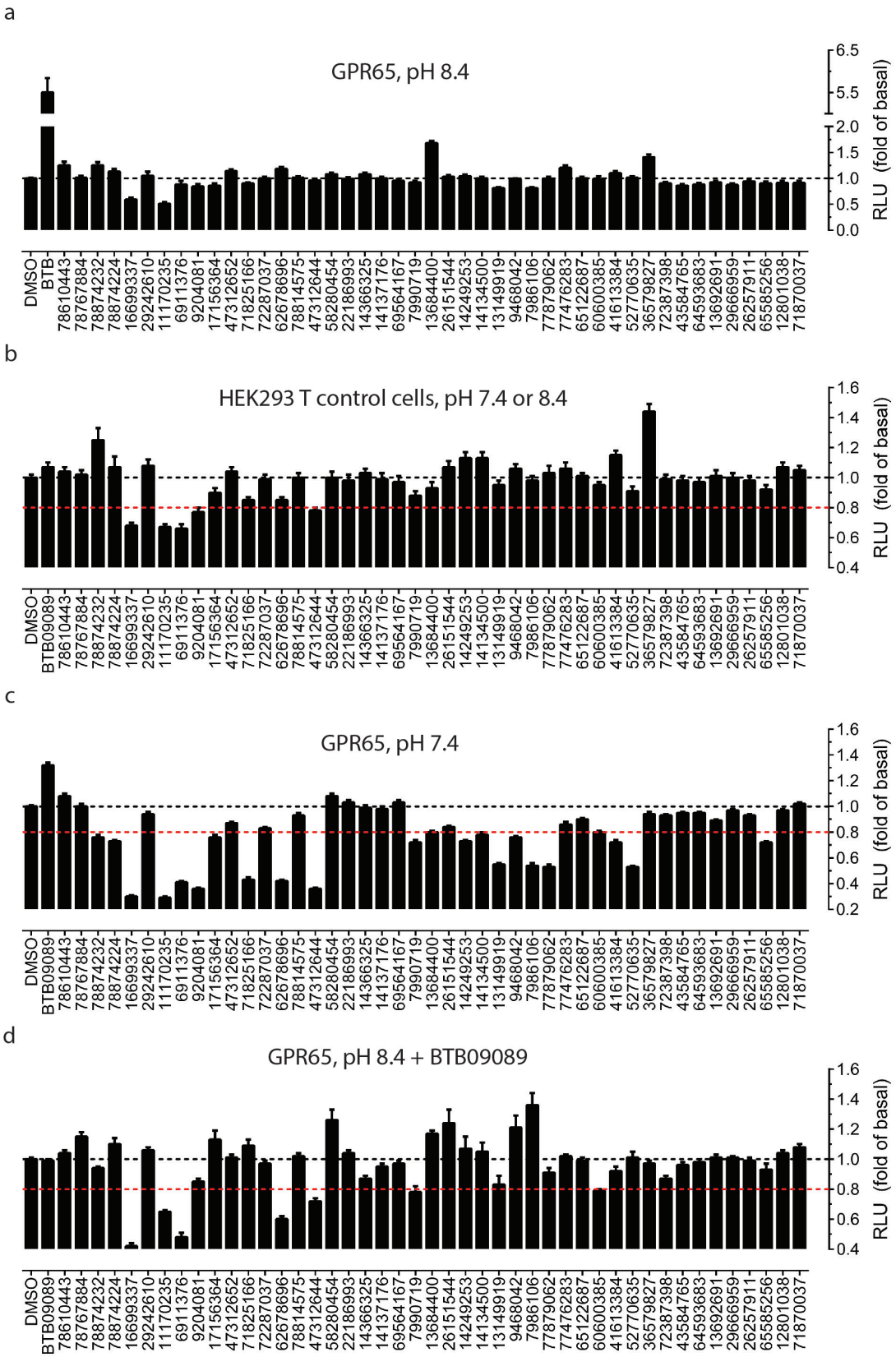




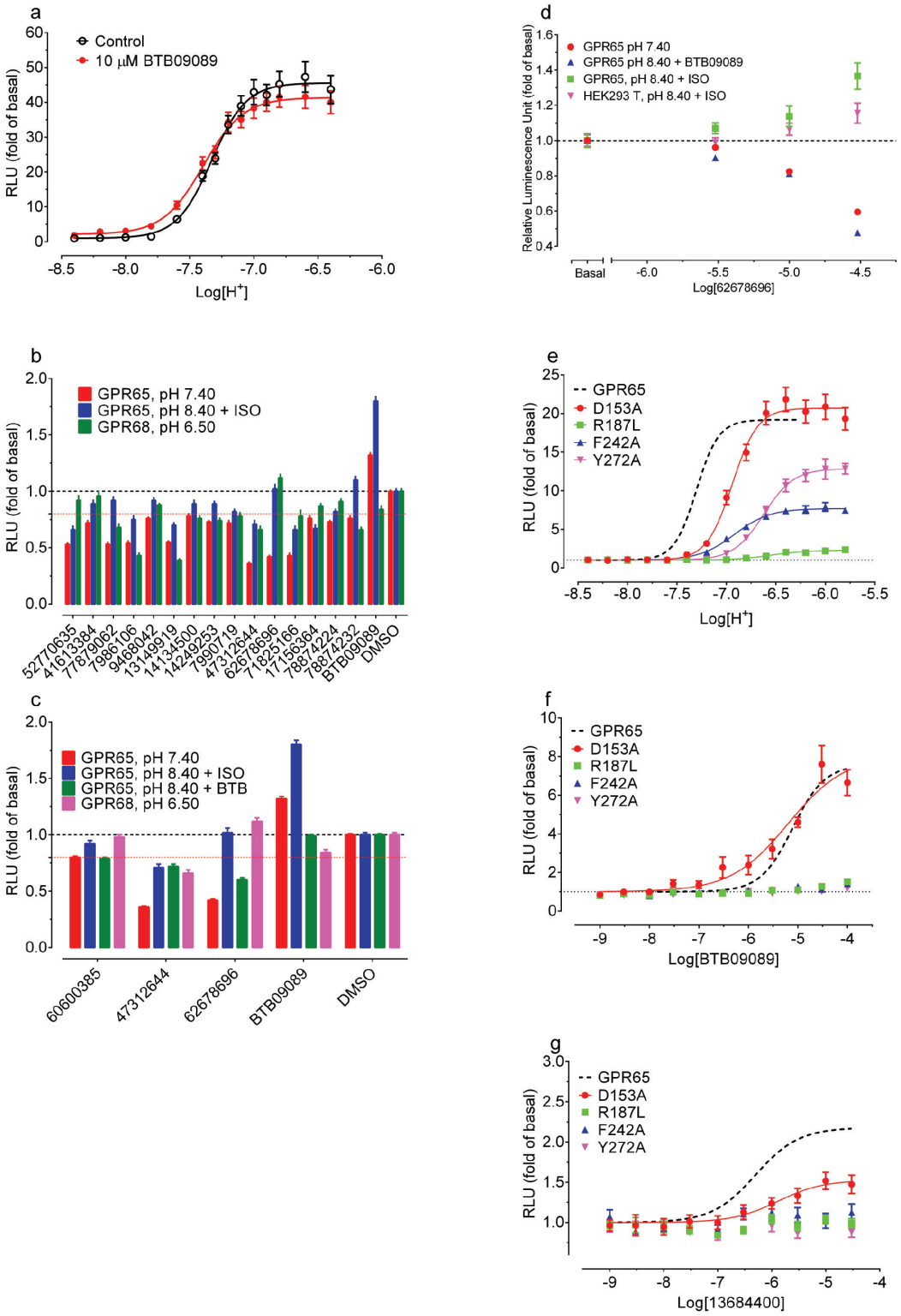
Extended Data Figure 7 | Characterization of potent GPR68 PAMs. **a**–**c**, Concentration–response curves of H^+ in the absence and presence of increasing concentrations of CGH2466 (**a**, **a′**), tracazolate (**b**, **b′**) and SLV320 (**c**, **c′**) and in the absence (left column, **a**, **b**, **c**) and presence (right column, **a′**, **b′**, **c′**) of phosphodiesterase inhibitor (Ro 20-1724, 30 μM) at GPR68-expressing cells. Normalized results (mean ± s.e.m., $n = 8$ for CGH2466; $n = 5$ for tracazolate; $n = 5$ for SLV320 for left column and $n = 3$ for right column) were analysed using a four-parameter logistic function and the standard allosteric operational model (not shown). Allosteric parameters in absence of Ro 20-1724 are summarized in Supplementary Table 8. For each pair of fittings, the proton potency value (negative logarithm of the half-maximum effective concentration (pEC_{50})) from the agonist concentration–response curve (right) in the absence of testing compound was used as the pK_A for the allosteric operational model (left). **d**, Schematic showing the shared pharmacology among $GABA_A$, adenosine GPCRs and GPR68 ligands. Molecules along each edge of the triangle have been shown to have activity at both targets, whereas tracazolate, in the middle, shows activity at all three.



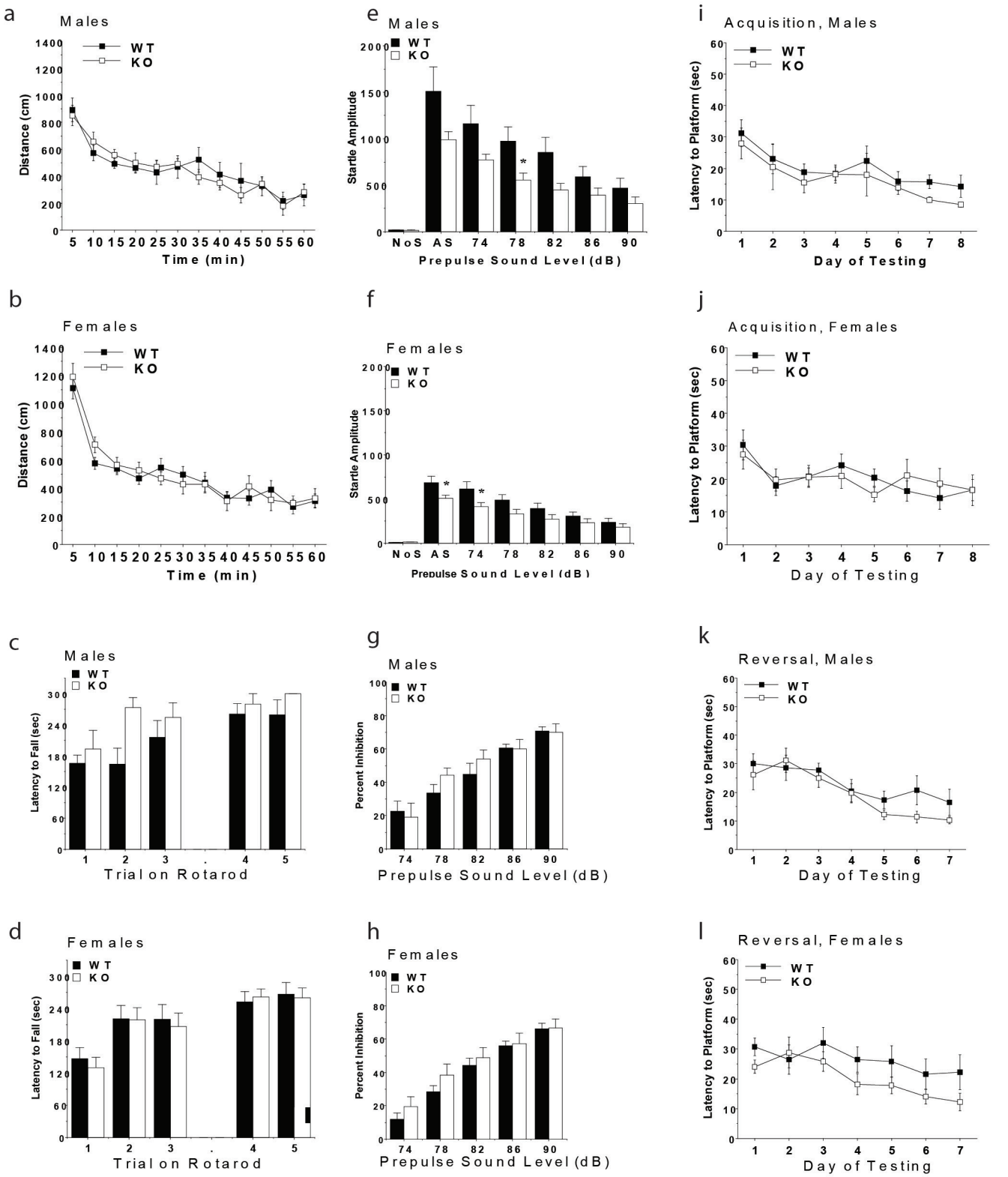
Extended Data Figure 8 | GPR68 mouse biology. a, Ogerin (Og) and lorazepam (Lo) activate PKA and p42/p44 MAP kinase in HEK293 cells stably expressing haemagglutinin (HA)-tagged GPR68 but not HApcDNA; vehicle (Ve). **b**, **c**, GPR68 knockout ($n = 7$) mice exhibited no differences in contextual memory retrieval (**b**) or cued memory retrieval (**c**) as compared to wild-type mice ($n = 8$). **d**–**f**, At 10 mg kg^{-1}, the ogerin isomer ZINC32547799 had no effect on learning (**d**) or contextual and cue memory (**e**, **f**) in wild-type mice (vehicle, $n = 6$; drug, $n = 7$). **g**–**i**, At 30 mg kg^{-1}, ZINC32547799 enhanced wild-type learning (**g**, drug × time interaction, $F_{(3.39)} = 3.58$, $P = 0.022$; drug alone $F_{(1.39)} = 1.19$, $P = 0.295$; Bonferroni post-hoc test revealed a significant effect ($P < 0.05$) at the third unconditioned/conditioned stimulus training point, two-way ANOVA), but had no effect at contextual and cue memory (**h**, **i**) (vehicle, $n = 7$; drug $n = 8$). Male mice at age of 6–8 weeks were used in the test. Normalized contextual memory retrieval (**d**) and cued memory retrieval (**f**) are presented in Fig. 4c, d.



Extended Data Figure 9 | Screening of ZINC compounds predicted to be active at GPR65 based on BTB09089 docking poses. a–d, Primary screening with ZINC compounds (30 μM) for agonist activity at GPR65 when receptors were kept inactive at pH 8.40 (**a**); at control HEK293T cells for nonspecific activity (**b**); at GPR65 when receptors were activated at pH7.40 for modulator or antagonist activity (**c**); at GPR65 when receptors were activated by BTB09089 (30 μM) at pH 8.40 for modulator or antagonist activity (**d**). Normalized results represent mean ± s.e.m. from a minimum of three assays (each in minimum of triplicate and a total of ≥16 measurements). The red dashed line in **b–d** indicates the 20% inhibition line (an arbitrary cut-off line).



Extended Data Figure 10 | Characterization of GPR65 allosteric modulators at wild-type and mutant receptors. a, BTB09089 showed weak agonist activity, but failed to potentiate proton activity at GPR65 ($n = 8$). **b**, Selected compounds from Extended Data Fig. 9b, c were tested for GPR65 specific inhibition ($n = 16$–56 measurements). Several compounds (such as ZINC41613384, ZINC9468042 and ZINC62678696) showed GPR65-specific inhibition. **c**, Selected compounds from Extended Data Fig. 9b, d were tested for antagonist activity against BTB09089-activated signal at GPR65 ($n = 16$–64 measurements). ZINC62678696 showed GPR65 specific inhibition when it was activated by either proton or BTB09089. **d**, ZINC62678696 inhibited GPR65 activity. **e**–**g**, Proton concentration–responses (**e**), BTB09089 concentration–responses (**f**), and ZINC13684400 concentration–responses (**g**) at GPR65 mutant receptors. Normalized results represent mean ± s.e.m. ($n \geq 3$) and curves were analysed in GraphPad Prism with a standard four-parameter logistic function. Corresponding curves of proton at GPR65 wild-type receptors (from Extended Data Fig. 4a) and BTB09089 and ZINC13684400 (from Fig. 5e) are also included (dashed lines) for comparison. Pharmacological parameters are listed in Supplementary Table 13.



Extended Data Figure 11 | *In vivo* behavioural profiling of GPR68-knockout mice. a, **b**, No effects of GPR68 deletion on distance travelled in an open field. Data represent mean ± s.e.m. for each group for a one-hour test session. **c**, **d**, No difference on latency to fall from an accelerating rotarod. Data represent mean ± s.e.m. for each group. **e**–**h**, Decreased startle responses in GPR68 knockout mice after presentation of acoustic stimuli (**e**, **f**). Data represent mean ± s.e.m. for each group. No effects of genotype were found for levels of prepulse inhibition (**g**, **h**). Data represent mean ± s.e.m. for each group ($^*P < 0.05$). **i**–**l**, No difference at acquisition and reversal learning in the Morris water maze. Data represent mean ± s.e.m. of four trials per day. Subject numbers were 9 wild-type and 7 knockout male mice, and 12 wild-type and 11 knockout female mice.

LETTER

doi:10.1038/nature15747

Extremely metal-poor stars from the cosmic dawn in the bulge of the Milky Way

L. M. Howes[1], A. R. Casey[2], M. Asplund[1], S. C. Keller[1], D. Yong[1], D. M. Nataf[1], R. Poleski[3,4], K. Lind[5], C. Kobayashi[1,6], C. I. Owen[1], M. Ness[7], M. S. Bessell[1], G. S. Da Costa[1], B. P. Schmidt[1], P. Tisserand[1,8], A. Udalski[3], M. K. Szymański[3], I. Soszyński[3], G. Pietrzyński[3,9], K. Ulaczyk[3,10], Ł. Wyrzykowski[3], P. Pietrukowicz[3], J. Skowron[3], S. Kozłowski[3] & P. Mróz[3]

The first stars are predicted to have formed within 200 million years after the Big Bang[1], initiating the cosmic dawn. A true first star has not yet been discovered, although stars[2–4] with tiny amounts of elements heavier than helium ('metals') have been found in the outer regions ('halo') of the Milky Way. The first stars and their immediate successors should, however, preferentially be found today in the central regions ('bulges') of galaxies, because they formed in the largest over-densities that grew gravitationally with time[5,6]. The Milky Way bulge underwent a rapid chemical enrichment during the first 1–2 billion years[7], leading to a dearth of early, metal-poor stars[8,9]. Here we report observations of extremely metal-poor stars in the Milky Way bulge, including one star with an iron abundance about 10,000 times lower than the solar value without noticeable carbon enhancement. We confirm that most of the metal-poor bulge stars are on tight orbits around the Galactic Centre, rather than being halo stars passing through the bulge, as expected for stars formed at redshifts greater than 15. Their chemical compositions are in general similar to typical halo stars of the same metallicity although intriguing differences exist, including lower abundances of carbon.

Stars with a low content of heavy elements have distinct spectral flux distributions, which are reflected in their colours. Using the photometric filter system on the SkyMapper telescope operated by the Australian National University, it is possible to identify metal-poor candidate stars[10] in the Galactic halo[4] and bulge[9]. We have observed ~14,000 bulge stars preselected from SkyMapper photometry using the AAOmega spectrograph on the Anglo-Australian Telescope (AAT), which enables the acquisition of 400 simultaneous stellar spectra over a 2-degree field of view. More than 500 stars with an iron abundance less than 1/100th of the solar value have been identified, making our survey the first to successfully target metal-poor stars in the Milky Way bulge. Twenty-three of these stars, targeted as the most metal-poor ones on the basis of the intermediate resolution spectra (Extended Data Table 1), were observed in June 2014 with the MIKE high-resolution spectrograph on the 6.5-m Magellan Clay telescope[11] to enable a comprehensive determination of their chemical compositions (Fig. 1).

The stars' effective temperatures were derived through fitting the observed hydrogen lines with theoretical spectra, while neutral and ionized iron lines provided measurements of the surface gravities and metallicities in the framework of 1D stellar atmosphere models[12] and non-equilibrium spectral line formation[13] (Extended Data Table 2). All 23 stars were found to have [Fe/H] ≤ -2.3, including nine stars with [Fe/H] < -3 (here $[A/B] = \log_{10}(N_A/N_B)_* - \log_{10}(N_A/N_B)_\odot$, where N_A/N_B refers to the number ratio of atoms of elements A and B in the star ($*$ subscript) and the Sun ($\odot$ subscript)). The most metal-poor star, SMSS J181609.62−333218.7, has [Fe/H] $= -3.94 \pm 0.16$. The abundances of an additional 22 elements were determined spectroscopically, including the α-elements Mg, Si, Ca, and Ti, and the neutron capture elements Y, Zr, and Ba (Extended Data Tables 3, 4, 5).

To confirm their bulge membership, the distances and orbits of the stars have been determined. Using the spectroscopic temperatures and surface gravities, and an assumed mass of $0.8M_\odot$, distances were inferred, which in nearly all cases are consistent with them being located within the bulge (Fig. 2). We have measured velocities for ten of our stars using observations taken by the OGLE-IV survey[14], from which orbits around the Galaxy have been determined in combination with their distances and velocities (Extended Data Table 6); the remaining stars fall outside the OGLE footprint while other sources of kinematic information are too uncertain to constrain the orbits sufficiently. Seven out of the ten stars with accurate kinematics are shown to have tightly bound orbits, placing them in the inner regions of the Milky Way (Fig. 2). In particular, using a cut-off radius of 3.43 kpc as the radius of the bulge component[15], the most metal-poor star SMSS J181609.62−333218.7 has an orbit entirely contained within the bulge. Only two out of the ten stars are on much larger orbits, being merely halo stars currently passing through the bulge region. Extending these numbers to the whole sample, we can expect ~14 of the 23 bulge stars analysed here to have orbits fully within the central regions of the Milky Way; with the imminent arrival of kinematic data from the Gaia satellite, accurate orbits for all of the bulge stars will be able to be determined.

The very first stars are predicted to have brought about the cosmic dawn by forming in the centres of the largest dark matter mini-haloes, which subsequently accreted material to become the inner regions of the largest galaxies[16]. The typical redshift of formation for stars in the bulge with [Fe/H] < -1 is $z \approx 10$, in contrast to $z \approx 5$ for halo stars. Of the stars with [Fe/H] < -3, approximately 15% are expected to have formed at $z > 15$ (refs 5, 6). Of the ten stars with accurate orbit information, half of them have binding energies $E_{tot} < -8 \times 10^{-4}$ km^2 s^{-2}, which is consistent with a formation redshift of $z > 15$ (ref. 5). Low binding energies imply that the stars have been in the Galactic potential well for some time and it is very unlikely they have been accreted from a recent dwarf spheroidal merger. Their low metallicities, orbits and binding energies make these stars prime candidates for being direct descendants of the very first stars, probing a cosmic epoch otherwise completely inaccessible currently. Direct age determinations of these ancient and extremely metal-poor bulge stars from comparison with stellar evolutionary tracks or radioactive U or Th dating are currently not possible, but asteroseismic ages could possibly be inferred with the extended Kepler mission or future satellites.

[1]Research School of Astronomy and Astrophysics, Australian National University, Australian Capital Territory 2601, Australia. [2]Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK. [3]Warsaw University Observatory, Aleje Ujazdowskie 4, 00-478 Warszawa, Poland. [4]Department of Astronomy, Ohio State University, 140 West 18th Avenue, Columbus, Ohio 43210, USA. [5]Department of Physics and Astronomy, Division of Astronomy and Space Physics, Uppsala University, Box 516, SE-751 20 Uppsala, Sweden. [6]School of Physics, Astronomy and Mathematics, Centre for Astrophysics Research, University of Hertfordshire, College Lane, Hatfield AL10 9AB, UK. [7]Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany. [8]Sorbonne Universités, UPMC Université Paris 6 et CNRS, UMR 7095, Institut d'Astrophysique de Paris, 98 bis Boulevard Arago, 75014 Paris, France. [9]Universidad de Concepción, Departamento de Astronomia, Casilla 160-C, Concepción, Chile. [10]Department of Physics, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, UK.

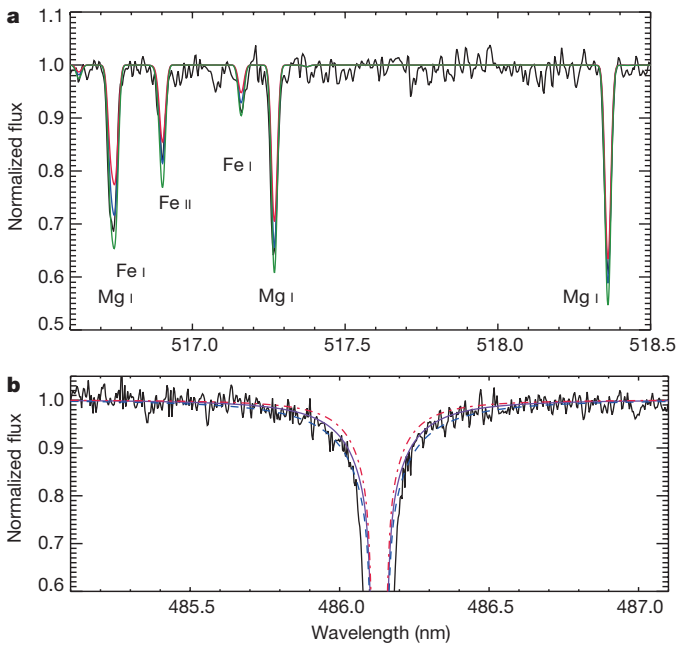


Figure 1 | Extracts of the spectrum of the lowest-metallicity star in our sample. a, A section of the spectrum of SMSS J181609.62−333218.7 (black line), the most metal-poor bulge star known. In blue is the predicted spectrum with the inferred stellar parameters (effective temperature $T_{\rm eff} = 4{,}809$ K, $\log(g) = 1.93$ (here surface gravity g is in cgs units), [Fe/H] = −3.94, [Mg/Fe] = 0.20), and the red and green lines show spectra with all abundances scaled to ± 0.15 dex, respectively. All three predicted spectra were created using the 1D local thermodynamic equilibrium (LTE) spectrum synthesis programme, MOOG[26]. **b**, The Hβ line of the same star, compared to three synthetic spectral line profiles[27] computed with $T_{\rm eff} = 4{,}640$ K (red, dash-dot), 4,800 K (purple, continuous), and 4,960 K (blue, dashed).

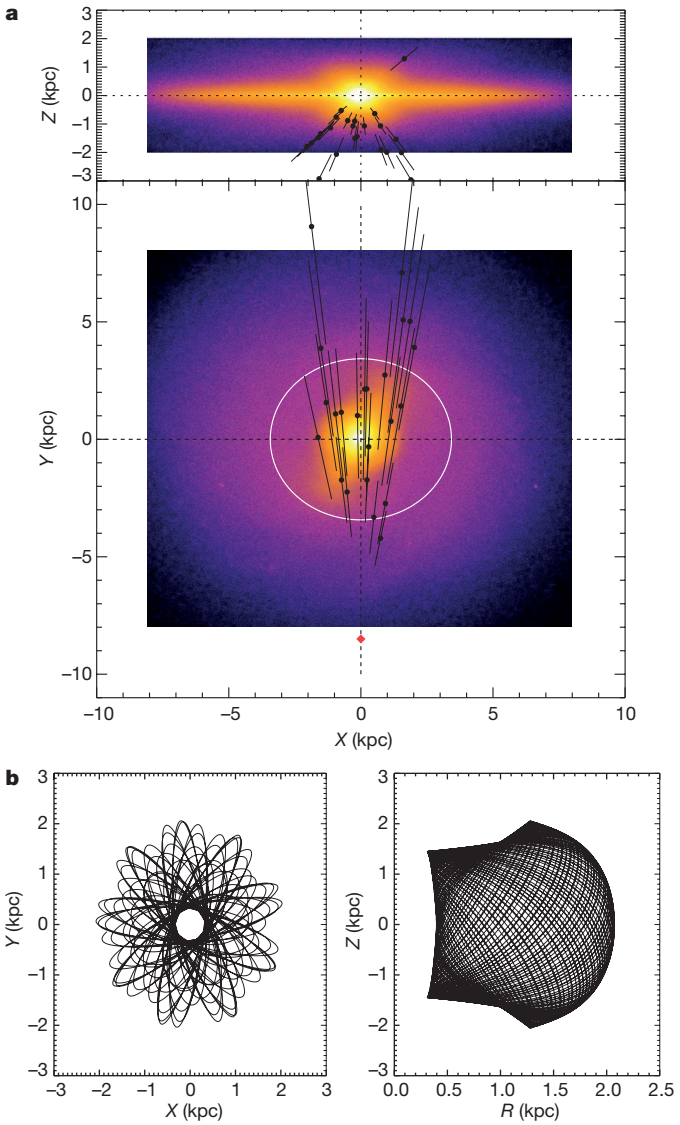


Figure 2 | The Galactic positions and orbits of the 23 stars observed at high resolution. a, Surface density map of a model of the Galactic bulge projected onto the X–Z (top) and X–Y (bottom) planes[28], where X, Y, and Z are Cartesian coordinates with the origin at the Galactic Centre and Z perpendicular to the plane of the Galaxy. Plotted over this (filled black circles) are the 23 stars of this study, with distance uncertainties shown as error bars, and a circle of radius 3.43 kpc (white: the cut-off radius of the inner bulge determined from 2MASS data[15]). The position of the Sun is shown with a red diamond, at 8.5 kpc from the Galactic Centre. **b**, Projections of the orbit of the lowest metallicity star, SMSS J181609.62−333218.7, both in the (R, Z) plane (right), where R is the radial direction, and in the plane of the orbit itself (left).

Given their extremely low metallicities and large formation redshifts, these stars are likely to have formed from gas polluted by ejecta from a single or at most a few supernovae of the first stellar generation. A chemical composition analysis has been carried out to search for tell-tale nucleosynthetic signatures and possible differences from halo stars at the same metallicities. For most elements, the chemical compositions of the 23 bulge stars are consistent with typical halo stars, suggesting enrichment by similar supernovae in spite of the distinct environments and formation redshifts. Subtle differences do exist however, most notably in terms of the carbon abundances. None of the 23 stars have the large observed carbon enhancements that occur frequently in halo stars. Applying evolutionary corrections to the surface carbon abundance to counter the mixing that occurs with material processed by H-burning through CNO-cycling at late stages of the stellar lifetime[17], still only one of the stars would have had a natal [C/Fe] > 1. In the halo, the percentage of stars that are carbon-enhanced increases dramatically at lower metallicities—from 27% of stars with [Fe/H] < −2 up to 69% with [Fe/H] < −4 (ref. 17). From the literature data on halo stars with similar iron abundances to our stars[17], the probability of selecting at most one carbon-enhanced star out of 23 halo stars is only 0.2%. Carbon-enhanced stars come in two varieties: those with and those without large excesses of neutron-capture elements. The former are most likely to have been formed by mass-transfer from a binary companion that underwent the asymptotic giant branch phase. Those carbon-enhanced stars with neutron-capture excesses occur most frequently at metallicities of [Fe/H] > −3, whereas those without do not appear to have binary companions, and are more common at the very lowest metallicities. As none of our bulge stars are classified as having large abundances of neutron capture elements, the likelihood of finding one such carbon-enhanced star out of 23 is 7% if the frequency is the same for the bulge as for the halo. A lower frequency of carbon-enhanced stars in the bulge relative to the halo is contrary to theoretical predictions; the expected dependence of the initial mass function on the cosmic microwave background[18] would result in a greater number of carbon-enhanced stars near the centre of the Galaxy.

The most metal-poor bulge star, SMSS J181609.62−333218.7, is at least an order of magnitude more iron-deficient than previously found low-metallicity bulge stars[9,19]. We have not been able to detect C in its spectrum, instead finding only an upper limit to its C abundance (Extended Data Fig. 1). This makes the upper limit on its total metallicity [Z/H] ≈ −3.8 (total mass fraction of $Z \approx 2.1 \times 10^{-6}$), where Z represents the sum of all metals, placing it amongst the four most metal-poor stars known, along with the halo star SDSS J102915+172927 (ref. 3). The low C measured in both these stars fall below the predicted

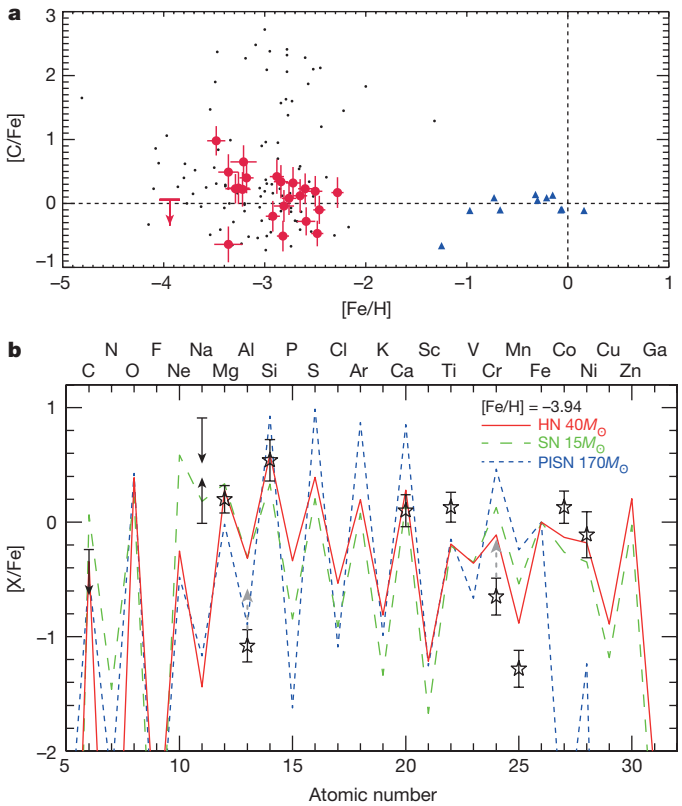


Figure 3 | Chemical abundances of the 23 stars observed at high resolution. a, The abundance ratio of carbon versus iron ([C/Fe]), with respect to metallicity ([Fe/H]) measured in the observed stars (filled red circles, red arrow for an upper limit). The dotted lines represent the solar abundances. Also shown for comparison are literature metal-poor halo giants (small black dots[24]) and more metal-rich bulge stars (filled blue triangles[29]). **b**, The chemical abundance pattern of SMSS J181609.62−333218.7, for elements X, where X is displayed at the top of the figure. Each determined abundance is shown as an open black star. These abundances are compared to three synthetic supernovae yields: a pair-instability supernova of 170$M_\odot$ (PISN; blue, short dash[22]), a core-collapse supernova of 15$M_\odot$ (SN: green, long dash[21]), and a hypernova of 40$M_\odot$ (HN; red, solid[21]). Dashed grey arrows represent expected non-LTE corrections; solid arrows represent measurements where only an upper or lower limit was possible. The error bars in **a** and **b** are estimates of the uncertainties in our measurements, calculated as described in Methods.

metallicity limit for formation of low-mass stars due to metal line cooling[20].

We have compared the detailed chemical abundance pattern of SMSS J181609.62−333218.7 to primordial supernovae yields[21,22] (Fig. 3). In particular, the low Mg and Ca abundance, but higher Si abundance, and the absence of a pronounced odd–even abundance pattern rule out the possibility of enrichment by a pair-instability supernova resulting from a primordial star of (140–250)$M_\odot$. Low abundances of Cr and Mn and of α-elements, combined with the higher abundance of Co, indicate that the polluting supernova was most likely to have been a primordial hypernova—an extremely energetic kind of supernova releasing ten times the kinetic energy of regular core-collapse supernovae, possibly due to the forming black hole having larger angular momentum[23]. Good agreement is found for a 40$M_\odot$ hypernova; a more stringent Zn limit would further constrain the mass range. Unusual abundance ratios have been found in small numbers of stars in the halo—4% of halo stars with low carbon abundances have chemical peculiarities in at least two elements[24]—but none so far appear to have been polluted by a 40$M_\odot$ hypernova. A low [α/Fe] ratio (0.14 dex; here α indicates α-elements) at such low [Fe/H] is consistent with an inhomogeneous enrichment from such supernovae[25], while stars with higher [α/Fe] formed from more well-mixed gas due to a longer time delay in forming the second generation of stars.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 1 May; accepted 24 August 2015.
Published online 11 November 2015.

1. Bromm, V., Yoshida, N., Hernquist, L. & McKee, C. F. The formation of the first stars and galaxies. *Nature* **459,** 49–54 (2009).
2. Christlieb, N. *et al.* A stellar relic from the early Milky Way. *Nature* **419,** 904–906 (2002).
3. Caffau, E. *et al.* An extremely primitive star in the Galactic halo. *Nature* **477,** 67–69 (2011).
4. Keller, S. C. *et al.* A single low-energy iron-poor supernova as the source of metals in the star SMSS J031300.36–670839.3. *Nature* **506,** 463–466 (2014).
5. Tumlinson, J. Chemical evolution in hierarchical models of cosmic structure. II. The formation of the Milky Way stellar halo and the distribution of the oldest stars. *Astrophys. J.* **708,** 1398–1418 (2010).
6. Salvadori, S., Ferrara, A., Schneider, R., Scannapieco, E. & Kawata, D. Mining the Galactic halo for very metal-poor stars. *Mon. Not. R. Astron. Soc.* **401,** L5–L9 (2010).
7. Feltzing, S. & Gilmore, G. Age and metallicity gradients in the Galactic bulge. *Astrophys. Space Sci.* **265,** 337–340 (1999).
8. García Pérez, A. E. *et al.* Very metal-poor stars in the outer Galactic bulge found by the APOGEE survey. *Astrophys. J.* **767,** L9 (2013).
9. Howes, L. M. *et al.* The Gaia-ESO survey: the most metal-poor stars in the Galactic bulge. *Mon. Not. R. Astron. Soc.* **445,** 4241–4246 (2014).
10. Keller, S. C. *et al.* The SkyMapper Telescope and the Southern Sky Survey. *Publ. Astron. Soc. Aust.* **24,** 1–12 (2007).
11. Bernstein, R., Shectman, S. A., Gunnels, S. M., Mochnacki, S. & Athey, A. E. MIKE: a double echelle spectrograph for the Magellan Telescopes at Las Campanas Observatory. *Proc. SPIE* **4841,** 1694–1704 (2003).
12. Gustafsson, B. *et al.* A grid of MARCS model atmospheres for late-type stars. *Astron. Astrophys.* **486,** 951–970 (2008).
13. Lind, K., Bergemann, M. & Asplund, M. Non-LTE line formation of Fe in late-type stars — II. 1D spectroscopic stellar parameters. *Mon. Not. R. Astron. Soc.* **427,** 50–60 (2012).
14. Udalski, A., Szymański, M. K. & Szymański, G. OGLE-IV: fourth phase of the Optical Gravitational Lensing Experiment. *Acta Astron.* **65,** 1–38 (2015).
15. Robin, A. C., Marshall, D. J., Schultheis, M. & Reylé, C. Stellar populations in the Milky Way bulge region: towards solving the Galactic bulge and bar shapes using 2MASS data. *Astron. Astrophys.* **538,** A106–A120 (2012).
16. Greif, T. H. *et al.* Formation and evolution of primordial protostellar systems. *Mon. Not. R. Astron. Soc.* **424,** 399–415 (2012).
17. Placco, V. M., Frebel, A., Beers, T. C. & Stancliffe, R. J. Carbon-enhanced metal-poor star frequencies in the Galaxy: corrections for the effect of evolutionary status on carbon abundances. *Astrophys. J.* **797,** 21 (2014).
18. Tumlinson, J. Carbon-enhanced metal-poor stars, the cosmic microwave background, and the stellar initial mass function in the early universe. *Astrophys. J.* **664,** L63–L66 (2007).
19. Schlaufman, K. C. & Casey, A. R. The best and brightest metal-poor stars. *Astrophys. J.* **797,** 13 (2014).
20. Frebel, A., Johnson, J. L. & Bromm, V. Probing the formation of the first low-mass stars with stellar archaeology. *Mon. Not. R. Astron. Soc.* **380,** L40–L44 (2007).
21. Kobayashi, C., Ishigaki, M. N., Tominaga, N. & Nomoto, K. The origin of low [α/Fe] ratios in extremely metal-poor stars. *Astrophys. J.* **785,** L5 (2014).
22. Umeda, H. & Nomoto, K. Nucleosynthesis of zinc and iron peak elements in population III type II supernovae: comparison with abundances of very metal poor halo stars. *Astrophys. J.* **565,** 385–404 (2002).
23. Nomoto, K. *et al.* Nucleosynthesis in black-hole-forming supernovae and extremely metal-poor stars. *Prog. Theor. Phys.* **151** (Suppl.), 44–53 (2003).
24. Yong, D. *et al.* The most metal-poor stars. II. Chemical abundances of 190 metal-poor stars including 10 new stars with [Fe/H] ≤ –3.5. *Astrophys. J.* **762,** 26–63 (2013).
25. Karlsson, T., Bromm, V. & Bland-Hawthorn, J. Pregalactic metal enrichment: the chemical signatures of the first stars. *Rev. Mod. Phys.* **85,** 809–848 (2013).
26. Sneden, C., Bean, J., Ivans, I., Lucatello, S. & Sobeck, J. MOOG: LTE line analysis and spectrum synthesis. *Astrophysics Source Code Library* http://adsabs.harvard.edu/abs/2012ascl.soft02009S (2012)
27. Barklem, P. S., Piskunov, N. & O'Mara, B. J. Self-broadening in Balmer line wing formation in stellar atmospheres. *Astron. Astrophys.* **363,** 1091–1105 (2000).
28. Ness, M. *et al.* Young stars in an old bulge: a natural outcome of internal evolution in the Milky Way. *Astrophys. J.* **787,** L19 (2014).
29. Ryde, N. *et al.* Chemical abundances of 11 bulge stars from high-resolution, near-IR spectra. *Astron. Astrophys.* **509,** A20–A35 (2010).

Acknowledgements This paper includes data gathered with the 6.5-m Magellan Telescopes located at Las Campanas Observatory, Chile. Australian access to the Magellan Telescopes was supported through the Collaborative Research Infrastructure Strategy of the Australian Federal Government. L.M.H. and M.A.

were supported by the Australian Research Council (FL110100012). A.R.C. acknowledges support from the European Union FP7 programme through ERC grant number 320360. Research on metal-poor stars with SkyMapper is supported through Australian Research Council Discovery Projects grants DP120101237 and DP150103294 (principal investigator G.S.D.C.). The OGLE project received funding from the NSC, Poland (MAESTRO grant 2014/14/A/ST9/00121 to A.U.).

Author Contributions The project was initiated and led by M.A. The photometric target selection was made by L.M.H., C.I.O. and D.M.N. using data from the SkyMapper telescope developed by B.P.S., S.C.K., G.S.D.C., M.S.B. and P.T. The low-resolution spectra were obtained by L.M.H and C.I.O. The data were reduced and analysed by L.M.H. using software developed by A.R.C. Target selection for the high-resolution observations was done by L.M.H., M.A. and A.R.C. with the observations carried out by L.M.H. and D.Y.; the reduction and subsequent chemical analysis was completed by L.M.H. K.L. performed the non-LTE spectral line formation calculations, C.K. interpreted the observed chemical abundances in terms of supernova yields, and M.N. provided comparison bulge data. R.P., A.U., M.K.S, I.S., G.P., K.U., Ł.W., P.P., J.S., S.K. and P.M. obtained the OGLE observations, A.U. and M.K.S. constructed the reference images, and R.P. measured the proper motions. The manuscript was written by M.A., L.M.H. and A.R.C. with all authors contributing comments.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to L.M.H. (louise.howes@anu.edu.au).

METHODS

Observations. Photometry of the Milky Way bulge was acquired for the EMBLA survey[9] during the commissioning period of the SkyMapper telescope in 2012 and 2013. Stars were selected from the photometry using a combination of the *g*, *i*, and *v* bandpasses, designed to give a reliable metallicity indicator[10].

Spectroscopic follow-up observations took place during 2012–14, making use of the AAOmega+2dF multi-object spectrograph[30] on the Anglo-Australian Telescope. With between 350 and 400 stars observed in each field, spectra of more than 14,000 bulge stars have been obtained. The gratings used have a spectral resolving power of 1,300 in the blue (370–580 nm) and 10,000 in the red (840–885 nm). The data were reduced using the standard 2dfdr pipeline. Stellar spectra were fitted using a generative model that simultaneously accounts for stellar parameters (by interpolating from the AMBRE grid[31]), continuum, spectral resolution and radial velocity.

From the first two years of spectroscopic data, more than 50 stars were identified as having [Fe/H] < -2.5. The high-resolution spectroscopic data of 23 stars presented in this Letter are the result of observations using the MIKE spectrograph at the Magellan Clay telescope[11] on 15–17 June 2014. All observations were taken using a slit width of 0.7″, resulting in a resolving power of 35,000 in the blue and 31,000 in the red. The data were reduced using the CarPy data reduction pipeline[32], before they were normalized and summed together using the SMH software[33]. The final spectra cover 330–890 nm.

Parameter and abundance determination. The stellar parameters (Extended Data Table 2) were calculated iteratively, using the original parameters from the low-resolution spectra as initial guesses. First, effective temperatures were derived by fitting the wings of the Balmer Hα and Hβ lines with a synthetic profile (Fig. 1). These profiles were created by linearly interpolating between a grid of synthetic spectra[27]. The best lines were fitted by a χ^2 minimization, using a weighted average of the two lines—weighting was double on the Hβ line, due to predicted LTE effects being larger for Hα (ref. 34). The difference between the temperatures calculated for each line was on average only 26 K. The log(*g*), microturbulence ξ_t, and [Fe/H] were then derived for that temperature, by forcing the Fe I abundance to remain constant with respect to reduced equivalent width, and equilibrium between the Fe I and Fe II abundances. Fe I and Fe II abundances were measured from the equivalent widths of a maximum of 66 Fe I lines and 24 Fe II lines (in the case of the most metal-poor star, SMSS J181609.62−333218.7, these numbers are reduced to 10 Fe I lines and 4 Fe II lines). Finally a non-LTE correction is applied to the Fe I abundance, calculated by taking the average of the line-by-line corrections[12]. This correction forces an offset between the Fe I and Fe II abundances, thus replacing the initial equilibrium. This process is repeated until the parameters converge on a solution. Throughout we use the 1D MARCS model atmospheres[13], and a shortened version of the Gaia-ESO line list, with extra lines supplemented from ref. 35 due to our wider wavelength coverage. The stellar abundances are referenced to the solar abundances of ref. 36. This analysis method was tested on seven halo stars from the literature[25], and the offsets found were $T_{\rm eff} = +28$ K, log(*g*) $= -0.2$, and [Fe/H] $= -0.08$ (literature values minus our values).

The abundances were measured using the equivalent widths of atomic lines (that were all on the linear part of the curve-of-growth), except in the case of C (measured from the C–H molecular bands at 431.3 nm and 432.3 nm) and Ba (synthesized in order to account for hyperfine splitting). Non-LTE corrections were calculated for Li (ref. 37), Na (ref. 38), Mg, and Ca, and applied to the individual line abundances. The literature halo abundances of Mg and Ca (ref. 25) shown in Fig. 3 have also had a NLTE correction applied, in order to ensure a fair comparison. 3σ upper limits were derived for some elements in those stars where the lines were too weak to be detected (Extended Data Table 3). The abundance offsets compared to the literature values averaged 0.10 ± 0.19 across those elements measured in common. Owing to wavelengths covered in the SkyMapper metallicity filter, it is possible that stars with extremely high C abundance appeared to be more metal-rich, and so were not selected. However, a similar study of metal-poor stars discovered in the halo with SkyMapper[39] found the fraction of C-enhanced stars was identical to that reported in previous surveys[40]. Furthermore, we followed up 14,000 stars with intermediate resolution spectra, and determined metallicities using those spectra. The majority of the stars observed had [Fe/H] ≈ -1.0 and included some that had solar metallicities, so it is highly unlikely that we missed any C-enhanced extremely metal-poor stars in our selection.

The systematic uncertainties in the temperature determinations were estimated to be ± 100 K, and the statistical uncertainties averaged ± 125 K, so when these are combined in quadrature we conclude the total uncertainty to be ± 160 K. The ξ_t uncertainties are estimated to be ± 0.2, mostly due to systematics.

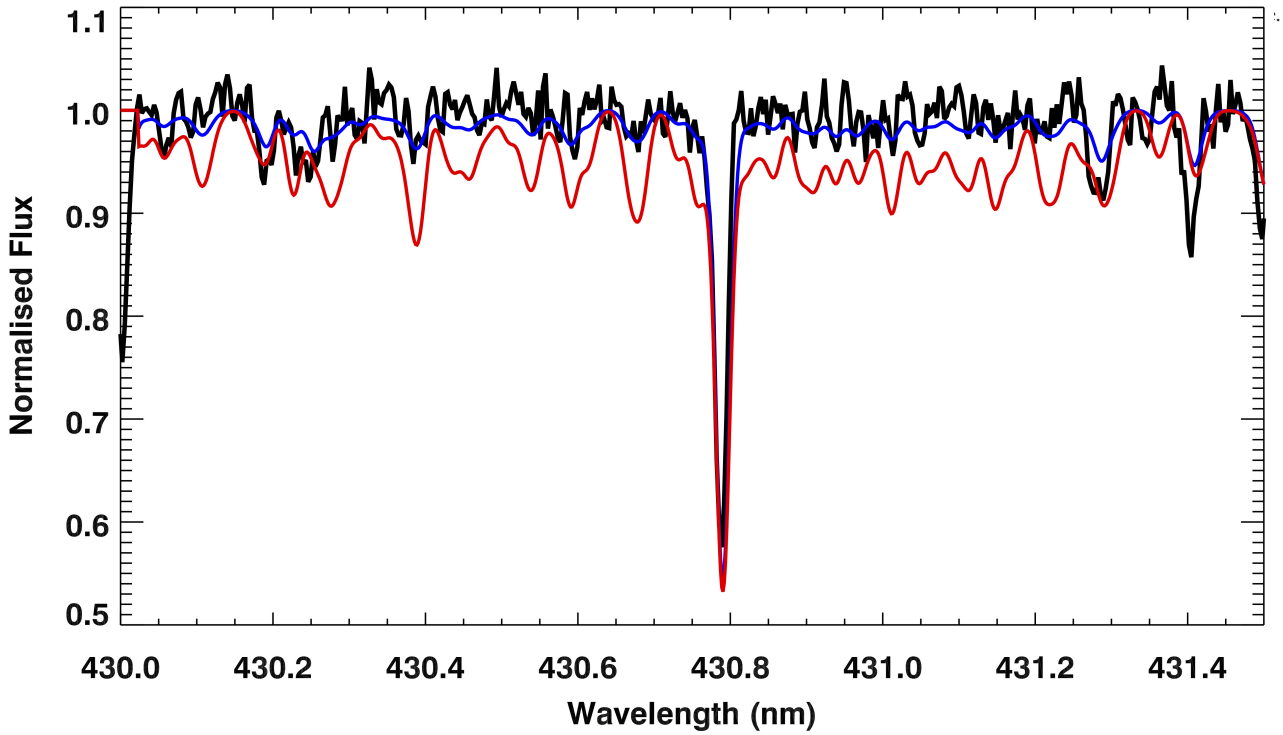
The standard errors of the individual line abundances of Fe I and Fe II were combined in quadrature to evaluate the log(*g*) uncertainties. The differences between the [Fe/H] values when varying the temperature, surface gravity, and microturbulence by their respective errors were combined in quadrature with the standard error of the Fe II lines to produce the [Fe/H] uncertainties. The individual abundance errors were also calculated using this method, using the standard error of the individual abundances for the lines of that particular element.

For SMSS J181609.62−333218.7, which has [Fe/H] $= -3.94$, a measurement of Na was not possible, owing to the 818.3 nm and 819.4 nm lines being too weak, and the Na D lines (588.9 nm and 589.5 nm) being partly blended with interstellar Na lines. We have derived a range of possible values for this star, taking the upper limit from the non-detection at 819.4 nm, and the lower limit from fitting a Gaussian to the Na D lines, taking into account the interstellar Na.

Distances and orbital parameters. Distances to the stars were calculated by comparing the absolute and apparent bolometric magnitudes. The absolute magnitudes were recovered from the relation $M_* = M_\odot - 2.5\log(L_*/L_\odot)$, where the luminosities are calculated using $L_*/L_\odot = 4\pi\sigma T^4 M_* G / 10^{\log(g_*)}$, taking $M_* = (0.8 \pm 0.2) M_\odot$ for all stars. The apparent bolometric magnitudes are reconstructed from the 2MASS JHK_s magnitudes (Extended Data Table 1), assuming reddening[41] (as no more-recent reddening catalogue covers all 23 stars), via the methodology of ref. 42. The proper motions are based on I band images taken during the OGLE-IV[14] observations of the Galactic bulge. Relative proper motions were derived from multiple epochs of data for each field[43], and the uncertainties are a combination of statistical and systematic (for each star, the systematic uncertainty is estimated to be ~0.4 mas yr^{-1}). These were converted into absolute proper motions by adding the predicted average bulge motion for each field, calculated using the Besancon Galaxy model[44]. The orbits were calculated using the python package galpy[45] and the galactic potential assumed in these calculations was a three-component Milky Way-like potential[46]. To model the uncertainty distributions, we sampled 1,000 orbits using a Monte Carlo simulation, assuming a normal distribution for the uncertainties of the input parameters. The results of this are included in Extended Data Table 6. One star, SMSS J175455.52 – 380339.3, has an unbound $E_{\rm tot}$ and impractically large orbital parameters, suggesting that one or more of our input parameters need to be changed.

Code availability. All codes used to analyse the data presented are publicly available. In particular, the 1D LTE analysis used was made possible with the line analysis and spectrum synthesis code MOOG[26].

30. Sharp, R. *et al.* Performance of AAOmega: the AAT multi-purpose fiberfed spectrograph. *Proc. SPIE* **6269,** 1–13 (2006).
31. de Laverny, P., Recio-Blanco, A., Worley, C. C. & Plez, B. The AMBRE project: a new synthetic grid of high-resolution FGKM stellar spectra. *Astron. Astrophys.* **544,** A126–A137 (2012).
32. Kelson, D. MIKE pipeline http://code.obs.carnegiescience.edu/mike (2014).
33. Casey, A. R. A tale of tidal tails in the Milky Way. Preprint at http://arXiv.org/abs/1405.5968 (2014).
34. Barklem, P. S. Non-LTE Balmer line formation in late-type spectra: effects of atomic processes involving hydrogen atoms. *Astron. Astrophys.* **466,** 327–337 (2007).
35. Norris, J. E. *et al.* The most metal-poor stars. I. Discovery, data, and atmospheric parameters. *Astrophys. J.* **762,** 25 (2013).
36. Asplund, M., Grevesse, N., Sauval, A. J. & Scott, P. The chemical composition of the Sun. *Annu. Rev. Astron. Astrophys.* **47,** 481–522 (2009).
37. Lind, K., Asplund, M. & Barklem, P. S. Departures from LTE for neutral Li in late-type stars. *Astron. Astrophys.* **503,** 541–544 (2009).
38. Lind, K., Asplund, M., Barklem, P. S. & Belyaev, A. K. Non-LTE calculations for neutral Na in late-type stars using improved atomic data. *Astron. Astrophys.* **528,** A103–A112 (2011).
39. Jacobson, H. R. *et al.* High-resolution spectroscopic study of extremely metal-poor star candidates from the SkyMapper survey. *Astrophys. J.* **807,** 171 (2015).
40. Lucatello, S. *et al.* The frequency of carbon-enhanced metal-poor stars in the Galaxy from the HERES sample. *Astrophys. J.* **652,** L37–L40 (2006).
41. Schlegel, D. J., Finkbeiner, D. P. & Davis, M. Maps of dust infrared emission for use in estimation of reddening and cosmic microwave background radiation foregrounds. *Astrophys. J.* **500,** 525–553 (1998).
42. Casagrande, L., Portinari, L. & Flynn, C. Accurate fundamental parameters for lower main-sequence stars. *Mon. Not. R. Astron. Soc.* **373,** 13–44 (2006).
43. Poleski, R. *et al.* An asymmetric streaming motion in the Galactic bulge X-shaped structure revealed by OGLE-III proper motions. *Astrophys. J.* **776,** 76 (2013).
44. Robin, A. C., Reylé, C., Derrière, S. & Picaud, S. A synthetic view on structure and evolution of the Milky Way. *Astron. Astrophys.* **409,** 523–540 (2003).
45. Bovey, J. *galpy* http://github.com/jobovy/galpy (2015).
46. Bovy, J. galpy: A python library for galactic dynamics. *Astrophys. J.* **216,** 29 (2015).



Extended Data Figure 1 | The C–H band of SMSS J181609.62 – 333218.7. The C–H band is used to derive an upper limit for C in our most metal-poor star, SMSS J181609.62 – 333218.7. Synthetic spectra with abundances of [C/Fe] = 0.06 (blue) and [C/Fe] = 0.56 (red) are shown for comparison.

Extended Data Table 1 | Coordinates and 2MASS photometry of the 23 stars observed

Name (SMSS)	RA (°)	Dec (°)	l (°)	b (°)	J (mag)	H (mag)	K_S (mag)
J173823.38-145701.1	264.597	-14.950	11.1	8.7	10.85	10.22	10.03
J182048.26-273329.2	275.201	-27.558	5.0	-6.1	12.94	12.42	12.25
J183744.90-280831.1	279.437	-28.142	6.2	-9.7	12.29	11.69	11.53
J183647.89-274333.1	279.200	-27.726	6.5	-9.3	10.68	10.03	9.77
J183812.72-270746.3	279.553	-27.130	7.1	-9.3	13.38	12.79	12.61
J183719.09-262725.0	279.330	-26.457	7.7	-8.9	12.79	12.19	12.03
J184201.19-302159.6	280.505	-30.367	4.5	-11.5	14.52	14.08	14.00
J184656.07-292351.5	281.734	-29.398	5.9	-12.0	13.12	12.61	12.51
J181406.68-313106.1	273.528	-31.518	0.8	-6.6	12.12	11.56	11.35
J181317.69-343801.9	273.324	-34.634	357.9	-7.9	13.09	12.55	12.50
J181219.68-343726.4	273.082	-34.624	357.9	-7.7	12.80	12.28	12.15
J181609.62-333218.7	274.040	-33.539	359.2	-7.9	13.39	12.84	12.71
J181634.60-340342.5	274.144	-34.062	358.8	-8.3	12.56	11.99	11.90
J175544.54-392700.9	268.936	-39.450	352.0	-7.1	13.71	13.19	13.09
J175455.52-380339.3	268.731	-38.061	353.1	-6.3	11.98	11.39	11.26
J175746.58-384750.0	269.444	-38.797	352.8	-7.2	13.09	12.60	12.51
J181736.59-391303.3	274.402	-39.218	354.2	-10.8	12.06	11.54	11.37
J181505.16-385514.9	273.772	-38.921	354.2	-10.2	13.63	13.15	13.09
J181921.64-381429.0	274.840	-38.241	355.2	-10.6	13.64	13.13	13.03
J175722.68-411731.8	269.345	-41.292	350.5	-8.3	13.85	13.25	13.21
J175021.86-414627.1	267.591	-41.774	349.4	-7.4	11.74	11.23	11.17
J175636.59-403545.9	269.152	-40.596	351.1	-7.9	12.86	12.29	12.19
J175433.19-411048.9	268.638	-41.180	350.4	-7.8	11.94	11.43	11.33

RA, right ascension; Dec., declination; *l* and *b*, Galactic longitude and latitude, respectively.

Extended Data Table 2 | Stellar parameters of the 23 stars observed

Name (SMSS)	V_{helio} (km s^{-1})	$d_{\odot}$ (kpc)	$T_{\rm eff}$ (K)	$\log g$ (cgs)	[Fe/H] (dex)	ξ_t (km s^{-1})	[α/Fe] (dex)
J173823.38-145701.1	46.1	8.5	4599	0.99	-3.36	2.30	0.12
J182048.26-273329.2	51.5	6.0	4949	2.22	-3.48	1.90	0.37
J183744.90-280831.1	-132.6	17.6	4597	0.98	-2.92	2.05	0.33
J183647.89-274333.1	-381.4	6.6	4649	1.17	-2.48	2.50	0.30
J183812.72-270746.3	155.3	12.3	4873	1.74	-3.22	1.81	-0.01
J183719.09-262725.0	-244.7	10.0	4791	1.64	-3.18	1.81	0.32
J184201.19-302159.6	171.8	9.6	5136	2.55	-2.84	1.96	0.30
J184656.07-292351.5	91.0	9.5	4857	1.93	-2.76	1.83	0.34
J181406.68-313106.1	4.9	9.3	4821	1.48	-2.82	1.96	0.22
J181317.69-343801.9	139.3	6.5	5015	2.25	-2.28	1.48	0.41
J181219.68-343726.4	-386.2	8.0	4873	1.94	-2.50	1.93	0.32
J181609.62-333218.7	27.4	10.4	4809	1.93	-3.94	1.60	0.14
J181634.60-340342.5	-170.3	10.5	4821	1.61	-2.46	1.79	0.06
J175544.54-392700.9	-279.6	13.5	4857	1.83	-2.65	1.60	0.32
J175455.52-380339.3	23.5	13.5	4714	1.10	-3.36	1.80	0.08
J175746.58-384750.0	-59.4	9.1	5064	1.96	-2.81	2.36	0.29
J181736.59-391303.3	-177.9	15.7	4612	1.05	-2.59	2.09	0.32
J181505.16-385514.9	202.1	5.0	4962	2.73	-3.29	2.10	0.35
J181921.64-381429.0	-97.7	11.2	4917	2.02	-2.72	1.94	0.30
J175722.68-411731.8	63.8	12.4	4894	1.97	-2.88	2.02	0.19
J175021.86-414627.1	181.4	4.1	5015	2.12	-2.60	1.55	0.30
J175636.59-403545.9	-28.8	9.8	4934	1.79	-3.21	1.96	0.20
J175433.19-411048.9	-229.3	5.6	4912	1.91	-3.26	1.94	0.35

Symbols: $V_{\rm helio}$, heliocentric velocity; $d_{\odot}$, distance from the Sun to the star; $T_{\rm eff}$, effective temperature; log(g), stellar surface gravity; ξ_t, microturbulence; [α/Fe] = ([Mg/Fe] + [Ca/Fe] + [Ti/Fe])/3. Average uncertainties: velocity, 1.0 km s^{-1}; distance, 3.0 kpc; temperature, 160 K; microturbulence, 0.2 dex; log(g), 0.14 dex; [Fe/H], 0.09 dex; [α/Fe], 0.13 dex.

Extended Data Table 3 | Chemical abundances measured for each star, C to Ca

Name (SMSS)	A(Li)	[C/Fe]	[Na/Fe]	[Mg/Fe]	[Al/Fe]	[Si/Fe]	[K/Fe]	[Ca/Fe]
J173823.38-145701.1		0.49	0.04	0.17	-0.78	0.27		0.12
J182048.26-273329.2		0.98	-0.28	0.54	-0.63	0.96		0.30
J183744.90-280831.1	0.16	-0.20	-0.28	0.44	-0.52	0.58	0.36	0.25
J183647.89-274333.1		-0.47	-0.24	0.33	-0.66	0.51		0.18
J183812.72-270746.3	0.93	0.22	-0.39	0.05	-1.23	0.14		0.03
J183719.09-262725.0		0.40	-0.19	0.47	-0.77	0.36	0.41	0.25
J184201.19-302159.6		0.34	-0.38	0.26	-0.89	0.38	0.53	0.37
J184656.07-292351.5	1.04	0.08	-0.30	0.41	-0.95	0.36	0.58	0.28
J181406.68-313106.1		-0.51	0.18	0.23	-0.94	0.32		0.16
J181317.69-343801.9	1.05	0.17	-0.33	0.53	-0.82	0.33	0.63	0.34
J181219.68-343726.4	1.01	0.19	-0.22	0.30	-0.86	0.25		0.31
J181609.62-333218.7		$<$0.06	-0.01$<$0.91[a]	0.20	-1.08	0.54		0.00
J181634.60-340342.5		-0.10	-0.53	0.05	-1.08	0.11	0.21	0.03
J175544.54-392700.9	0.87	0.12	-0.32	0.29	-0.88	0.34	0.36	0.29
J175455.52-380339.3		-0.64		0.06	-0.88	0.30		0.03
J175746.58-384750.0		-0.04		0.37	-1.10	0.44		0.24
J181736.59-391303.3		-0.28	-0.11	0.38	-0.69	0.53	0.53	0.26
J181505.16-385514.9		0.23	-0.23	0.21	-0.96	0.15		0.36
J181921.64-381429.0	1.04	0.32	-0.24	0.28	-0.82	0.54	0.44	0.26
J175722.68-411731.8		0.42	-0.42	0.21	-0.70	0.48	0.14	0.13
J175021.86-414627.1	0.98	0.23	-0.37	0.30	-0.82	0.42		0.28
J175636.59-403545.9	0.93	0.65		0.30	-0.76	0.45		0.11
J175433.19-411048.9	0.92	0.24	-0.03	0.40	-0.74	0.43	0.44	0.32

A(Li) is the logarithmic abundance of lithium. All abundances are derived using LTE, except for Li, Na, Mg, and Ca, where non-LTE corrections have been applied. Average uncertainties: Li, 0.20; C, 0.25; Na, 0.20; Mg, 0.16; Al, 0.22; Si, 0.21; K, 0.17; Ca, 0.12.

[a]−0.01 is the lower limit, and 0.91 is the upper limit; see Methods for details.

Extended Data Table 4 | Chemical abundances measured for each star, Sc to Cu

Name (SMSS)	[Sc/Fe]	[Ti/Fe]	[Cr/Fe]	[Mn/Fe]	[Co/Fe]	[Ni/Fe]	[Cu/Fe]
J173823.38-145701.1		-0.09	-0.22	-0.80	0.17	-0.21	<0.96
J182048.26-273329.2		0.16	-0.51	-0.97	0.24	-0.33	<1.33
J183744.90-280831.1	0.04	0.20	-0.23	-0.38	0.36	0.14	<0.29
J183647.89-274333.1	0.14	0.34	-0.27	-0.35	0.01	0.02	-0.43
J183812.72-270746.3		-0.20	-0.51	-0.32	0.22	-0.08	<1.06
J183719.09-262725.0	0.18	0.14	-0.33	-0.34	0.23	0.23	<1.10
J184201.19-302159.6	-0.03	0.20	-0.24	-0.57	0.35	-0.02	<0.72
J184656.07-292351.5	0.11	0.28	-0.19	-0.31	0.11	0.07	<0.45
J181406.68-313106.1	0.08	0.19	-0.30	-0.60	0.22	0.09	<0.50
J181317.69-343801.9	0.11	0.34	-0.19	-0.08	0.12	0.10	<0.13
J181219.68-343726.4	0.18	0.31	-0.15	-0.14	0.22	0.22	<0.17
J181609.62-333218.7		0.13	-0.65	-1.28	0.13	-0.11	<1.57
J181634.60-340342.5	-0.26	0.04	-0.24	-0.35	-0.20	-0.03	<-0.05
J175544.54-392700.9	-0.05	0.32	-0.22	-0.18	0.24	-0.05	<0.32
J175455.52-380339.3		0.02	-0.47	-1.05	0.07	-0.12	<0.85
J175746.58-384750.0	0.12	0.19	-0.44	-0.59	0.09	-0.01	<0.79
J181736.59-391303.3	0.15	0.24	-0.31	-0.34	0.00	0.08	<0.10
J181505.16-385514.9	0.40	0.38	-0.54	-0.81	0.17	-0.11	<1.02
J181921.64-381429.0	0.01	0.32	-0.17	-0.44	0.33	0.25	<0.50
J175722.68-411731.8	-0.22	0.15	-0.31	-0.48	-0.04	0.04	<0.82
J175021.86-414627.1	-0.23	0.26	-0.27	-0.31	0.16	0.16	<0.27
J175636.59-403545.9	-0.28	0.06	-0.29	-0.72	0.17	-0.26	<0.95
J175433.19-411048.9		0.22	-0.52	-0.91	0.29	0.02	<1.08

All abundances are derived using LTE. Average uncertainties: Sc, 0.10; Ti, 0.10; Cr, 0.21; Mn, 0.25; Co, 0.23; Ni, 0.19; Cu, 0.25.

Extended Data Table 5 | Chemical abundances measured for each star, Zn to Eu

Name (SMSS)	[Zn/Fe]	[Sr/Fe]	[Y/Fe]	[Zr/Fe]	[Ba/Fe]	[La/Fe]	[Eu/Fe]
J173823.38-145701.1	0.66	0.03	0.02	0.23	-0.04	-0.10	
J182048.26-273329.2	<1.01	-0.47			0.03	<1.33	
J183744.90-280831.1	0.27	-0.29	-0.32	0.03	-0.31	<0.12	
J183647.89-274333.1	0.23	0.18	-0.20	0.45	0.13	0.17	0.82
J183812.72-270746.3	<0.79	-1.03			-0.70	<0.77	
J183719.09-262725.0	0.48	0.04	0.53		-0.51	<1.03	
J184201.19-302159.6	<1.15	-0.20	0.04	0.70	0.16	<0.94	
J184656.07-292351.5	0.48	-0.26	-0.51	0.14	-0.32	<0.36	
J181406.68-313106.1	0.42	-1.61			-0.72	<0.32	
J181317.69-343801.9	0.17	0.17	-0.11	0.34	0.22	-0.09	0.15
J181219.68-343726.4	0.33	-0.06	-0.11	0.09	0.13	<0.90	0.48
J181609.62-333218.7	<1.40	-0.85	0.23		<-0.66	<1.09	0.91
J181634.60-340342.5	0.21	-0.25	-0.65	-0.21	-0.32	-0.14	-0.11
J175544.54-392700.9	0.36	-0.10	-0.2	0.38	-0.11	-0.15	0.21
J175455.52-380339.3	0.63	0.47	0.01	0.14	-0.57	<0.66	
J175746.58-384750.0		-0.21	0.04	0.91	0.23	<1.26	0.65
J181736.59-391303.3	0.23	-0.14	-0.47	0.14	-0.28	<1.19	0.21
J181505.16-385514.9	<0.95	-0.19	0.14	0.71	0.04	<0.54	0.96
J181921.64-381429.0	0.42	-0.21	-0.14	0.51	-0.01	0.48	0.59
J175722.68-411731.8	<0.95	-0.30	-0.30	0.24	-0.19	<0.63	0.52
J175021.86-414627.1	0.41	-0.14	-0.40	0.25	-0.08	<0.60	0.23
J175636.59-403545.9	0.86	0.55	0.24	0.71	-0.95	<1.21	
J175433.19-411048.9	0.50	-0.81	-0.29		-0.42	<0.91	

All abundances are derived using LTE. Average uncertainties: Zn, 0.10; Sr, 0.20; Y, 0.12; Zr, 0.12; Ba, 0.17; La, 0.15; Eu, 0.16.

Extended Data Table 6 | Orbital parameters

Name (SMSS)	$\mu_\alpha \cos\delta$ (mas yr^{-1})	μ_δ (mas yr^{-1})	Mean r_{peri} (kpc)	Mean r_{ap} (kpc)	Mean Eccentricity	Mean Z_{max} (kpc)	E_{tot} (10^4 km^2 s^2)
J182048.26-273329.2	-4.10 $\pm$0.52	-6.38 $\pm$0.51	0.5 $^{+0.9}_{-0.2}$	2.9 $^{+1.7}_{-1.1}$	0.70 $^{+0.11}_{-0.23}$	0.9 $^{+0.2}_{-0.5}$	-11.3 $^{+2.2}_{-1.4}$
J184201.19-302159.6	-0.38 $\pm$0.90	-0.82 $\pm$0.90	1.2 $^{+1.1}_{-0.7}$	6.6 $^{+4.3}_{-1.6}$	0.72 $^{+0.13}_{-0.20}$	4.3 $^{+1.7}_{-1.8}$	-7.1 $^{+2.1}_{-1.2}$
J184656.07-292351.5	1.17 $\pm$0.89	-2.32 $\pm$0.89	1.2 $^{+1.5}_{-0.8}$	4.8 $^{+2.7}_{-1.4}$	0.65 $^{+0.19}_{-0.41}$	3.2 $^{+1.2}_{-1.4}$	-8.3 $^{+1.9}_{-1.2}$
J181406.68-313106.1	2.28 $\pm$0.52	-8.25 $\pm$0.52	1.1 $^{+2.3}_{-0.8}$	3.5 $^{+9.1}_{-1.3}$	0.63 $^{+0.18}_{-0.16}$	2.7 $^{+8.6}_{-0.8}$	-9.5 $^{+5.4}_{-1.8}$
J181219.68-343726.4	-2.42 $\pm$1.14	-1.29 $\pm$1.14	0.7 $^{+0.8}_{-0.4}$	17.3 $^{+11.3}_{-4.9}$	0.92 $^{+0.05}_{-0.03}$	7.2 $^{+3.2}_{-1.8}$	-3.4 $^{+1.9}_{-1.4}$
J181609.62-333218.7	-4.14 $\pm$0.64	-3.74 $\pm$0.64	1.0 $^{+3.5}_{-0.8}$	3.4 $^{+3.3}_{-1.8}$	0.53 $^{+0.23}_{-0.28}$	1.9 $^{+2.6}_{-0.9}$	-9.9 $^{+4.0}_{-2.6}$
J181634.60-340342.5	1.92 $\pm$0.62	-0.31 $\pm$0.62	1.9 $^{+2.0}_{-1.3}$	7.9 $^{+8.3}_{-2.4}$	0.65 $^{+0.15}_{-0.10}$	3.9 $^{+4.2}_{-1.7}$	-6.3 $^{+3.0}_{-2.0}$
J175544.54-392700.9	0.03 $\pm$1.49	-0.35 $\pm$1.46	1.7 $^{+2.3}_{-1.2}$	29.3 $^{+67.4}_{-16.5}$	0.92 $^{+0.05}_{-0.04}$	10.8 $^{+29.6}_{-6.0}$	-1.4 $^{+3.7}_{-3.2}$
J175455.52-380339.3	1.98 $\pm$1.14	4.76 $\pm$1.14	5.5 $^{+4.8}_{-3.9}$	811.5 $^{+1216.1}_{-796.2}$	0.98 $^{+0.01}_{-0.16}$	81.0 $^{+259.4}_{-76.6}$	6.9 $^{+12.2}_{-10.7}$
J175746.58-384750.0	1.86 $\pm$1.25	0.17 $\pm$1.25	1.8 $^{+1.9}_{-1.2}$	5.9 $^{+5.0}_{-1.9}$	0.59 $^{+0.24}_{-0.28}$	2.3 $^{+2.8}_{-1.1}$	-7.5 $^{+2.9}_{-1.8}$

Symbols: $\mu_\alpha \cos\delta$ and μ_δ are the proper motions in equatorial coordinates; r_{peri} and r_{ap} are the pericentric and apocentric radii of the orbit, respectively, Z_{max} is the maximum distance the orbit reaches above/below the Galactic plane, and E_{tot} is the total energy of the orbit. All values given here are the mean values from the Monte Carlo simulation of 1,000 orbits.

LETTER

doi:10.1038/nature15731

Ubiquitous time variability of integrated stellar populations

Charlie Conroy[1], Pieter G. van Dokkum[2] & Jieun Choi[1]

Long-period variable stars arise in the final stages of the asymptotic giant branch phase of stellar evolution. They have periods of up to about 1,000 days and amplitudes that can exceed a factor of three in the I-band flux. These stars pulsate predominantly in their fundamental mode[1–3], which is a function of mass and radius, and so the pulsation periods are sensitive to the age of the underlying stellar population[4]. The overall number of long-period variables in a population is directly related to their lifetimes, which is difficult to predict from first principles because of uncertainties associated with stellar mass-loss and convective mixing. The time variability of these stars has not previously been taken into account when modelling the spectral energy distributions of galaxies. Here we construct time-dependent stellar population models that include the effects of long-period variable stars, and report the ubiquitous detection of this expected 'pixel shimmer' in the massive metal-rich galaxy M87. The pixel light curves display a variety of behaviours. The observed variation of 0.1 to 1 per cent is very well matched to the predictions of our models. The data provide a strong constraint on the properties of variable stars in an old and metal-rich stellar population, and we infer that the lifetime of long-period variables in M87 is shorter by approximately 30 per cent compared to predictions from the latest stellar evolution models.

In typical massive galaxies with $\sim 10^{11}$ stars, the variation in the total light due to long-period variables will be small, as the summed light curves of many such stars effectively cancel each other out (with random phases, the net effect scales as $N^{-1/2}$, where N is the number of stars). If the light is spread out over many (for example, $\sim 10^4$–10^6) pixels, then the number of stars per pixel can range from $\sim 10^4$ to 10^7 and in this regime the number of asymptotic giant branch stars per pixel is small and governed by Poisson statistics. This 'semi-resolved' regime is well known[5], and the expected surface brightness fluctuations due to Poisson statistics of rare luminous stars have been observed and studied in several hundred nearby galaxies[6–8]. We expect in this regime to be able to detect the presence of variable stars through the time dependence of the pixel flux (that is, the pixel light curve): essentially, every pixel is expected to 'shimmer' on time-scales of several hundred days.

In order to quantify the expected pixel shimmer, we created a stellar population model at solar metallicity that included the time-dependent effect of long-period variables. We started with a new library of stellar isochrones (J.C. *et al.*, submitted) that densely samples fast phases of stellar evolution, and assigned periods to evolved stars assuming that they pulsate in the fundamental mode[4]. We then used observations of variable stars in the Galactic bulge from the OGLE survey to estimate a period–amplitude relation in the I band[9–11]. A smooth surface brightness model of the giant elliptical galaxy M87 was used to specify the luminosity within pixels of size $0.2'' \times 0.2''$. The pixel luminosity was used to normalize the weights in the isochrone assuming a Salpeter initial mass function[12]. For each pixel the number of giants was drawn from a Poisson distribution, and the time evolution of the flux for each giant was given by its associated period and amplitude and initialized with a random phase. An illustration of this time-dependent model for M87 is shown in Fig. 1. The variable-star part of our model has a tunable parameter, the long-period variable star weight, which can be interpreted as the typical lifetime of such stars. Further details regarding the modelling are provided in Methods.

We sampled the model with the same cadence and applied the same photon counting uncertainties as used for existing observations of M87 (see below). The resulting model pixel light curves are shown in Fig. 2 (blue lines). These pixels were selected to have peak-to-peak flux variation $>1.5\%$. While rising and falling curves are clearly seen, one also sees that a <100-d observing window can by chance sample a light curve at a phase that appears relatively flat. A >200-d observing cadence would clearly be ideal for observing the effects of long-period variables in the integrated light of nearby galaxies.

To test these expected variations, we analysed archival data of the galaxy M87 from the Hubble Space Telescope (HST) collected over 72 d in 2005[13–15]. Imaging was obtained in both the F606W and F814W filters with the Advanced Camera for Surveys. We focused our analysis on the F814W imaging as the F606W data were generally of lower quality (owing both to a shorter exposure time and the fact that only a single exposure was obtained per visit, which made it difficult to clean the images of blemishes such as hot pixels and cosmic rays). The data were processed via the standard HST pipeline. In total 52 separate images, each with a depth of 1,440 s, were considered in this analysis. Globular clusters in the field were used to refine the astrometric alignment with subpixel shifts. Accurate subtraction of the background was achieved with several additional corrections to the standard HST pipeline, as detailed in Methods. Pixels that deviated by more than

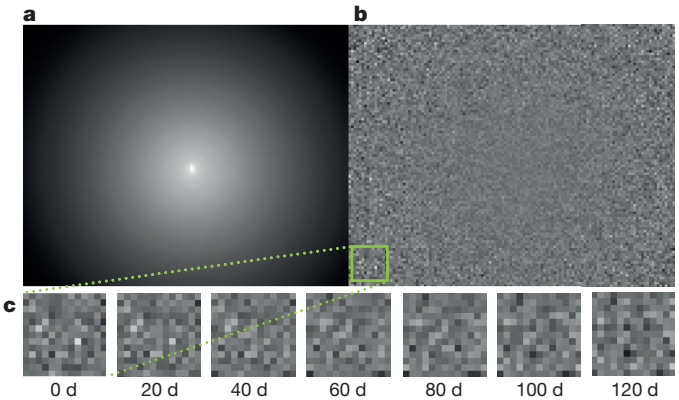


Figure 1 | Illustration of pixel shimmer. Model prediction of the effect of long-period variables on integrated light. **a**, A smooth model for the surface brightness profile of M87. **b**, The flux at time $t = 0$ divided by the mean flux over 1,000 d within each pixel. **c**, Zoom-in on the lower left corner (boxed in **b**), showing snapshots at 20-d intervals. Notice the coherent variation in brightness of individual pixels.

[1]Department of Astronomy, Harvard University, Cambridge, Massachusetts, USA. [2]Astronomy Department, Yale University, New Haven, Connecticut, USA.

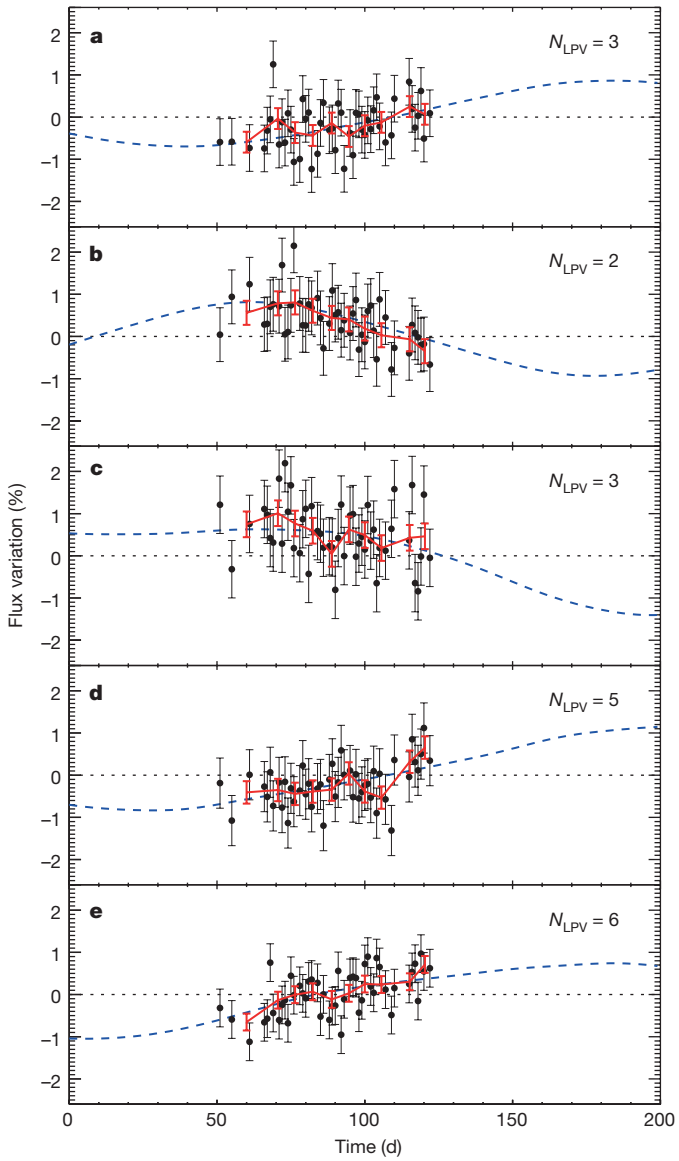


Figure 2 | Simulation of pixel light curves. a–e, Each panel shows modelled relative flux variation over 200 d for a randomly chosen pixel selected to have a peak-to-peak flux variation of >1.5%. The underlying stellar population is old and metal-rich, and includes long-period variable stars. In each panel, the noise-free model (dashed blue line) is compared to a simulation of the M87 data, including the photon counting noise (1σ errors) and cadence of the observations (filled black circles and error bars), and a boxcar average of the simulated data (red lines and 1σ error bars). Also shown in each panel is N_{LPV}, the number of long-period variable stars per pixel with periods >150 d.

30% from a smooth model of the light profile were masked. This effectively removed all visible globular clusters, background galaxies, the chip gap, and edge effects. We also masked the central region and the well-known jet in M87. The data were binned 4×4 to $0.2'' \times 0.2''$ in order to reduce the spatial coherency imposed by the point-spread function (PSF; note that the models were also spatially binned and have been convolved by the PSF in order to emulate the observations as closely as possible).

Example pixel light curves for M87 are shown in Fig. 3. The error bars represent photon counting uncertainties only; the solid line is a 5-point boxcar averaging of the data. We detect coherent variation in the pixel light curves that is qualitatively consistent with our model expectations. These examples were chosen to highlight the level of variety seen in the data. We note that a significant fraction of pixel light curves show no evidence of variation within the noise limits of the data. We show below that this is expected if long-period variables are the source of the variation.

We have quantified the pixel light curves by fitting each curve with a linear function; the best-fit slope and uncertainty were recorded. The resulting distribution of slopes (in units of the uncertainty) is shown in Fig. 4. We find that 24% of pixels (48,100 out of 202,000) show $>2\sigma$ evidence for variation. In our model there are on average 1.5 long-period variable stars responsible for variation in each $>2\sigma$ detection. This implies a statistical detection of ~72,000 variable stars in M87. When averaged over the central $1' \times 1'$ field of view, the model predicts on average 0.5 variable stars per pixel. In Fig. 4 we compare the observations to the model predictions. We show the sensitivity of the model light curve statistics to both the stellar population age and variable star parameters, and also the posterior probability distributions that result from fitting the model to the observed histogram when allowing the age and relative variable star weight to vary (we do not include the tails of the distribution in the fit as the data are slightly asymmetric beyond $|\text{slope/error}| \approx 7$). The variable star weight is an overall factor controlling the contribution of variable stars to the integrated light relative to the predictions of a stellar evolution model (see Methods for details). The pixel light curves provide a strong constraint on a combination of the age and variable star weight. The dashed line in Fig. 4d shows the best-fit age estimated from modelling the integrated light spectrum of the central region of M87[16], which allows us to break the degeneracy between age and variable star weight. It is noteworthy that the best-fit long period variable star weight is less than one, suggesting that such stars in M87 may have shorter lifetimes than current solar-metallicity stellar evolution models predict (see Methods for a discussion of the effects of metallicity).

This is not the first detection of time variability in the pixel fluxes of nearby galaxies; previous work predicted the occurrence of a gravitational microlensing signal at the pixel level[17], which was subsequently observed[18]. Novae have also been observed in nearby galaxies[19], and indeed we identified ~15 novae through visual inspection of the pixel light curves for M87. However, novae and microlensing events are rare (though bright) events. An important distinguishing feature of the time variation caused by long-period variable stars is the ubiquity—as Fig. 4 shows, 24% of the pixels show $>2\sigma$ evidence for variation.

There are relatively few constraints on the stellar evolutionary phase that gives rise to long-period variables. The best constraints to date on this phase are confined to the Magellanic Clouds, which have sub-solar metallicities characteristic of low-mass galaxies. The observations reported here have provided a direct constraint on this important stellar evolutionary phase in a massive, high-metallicity galaxy. New stellar evolution models over-predict the lifetimes of long-period variables by approximately 30% if a spectroscopic age for M87 of 10 Gyr is adopted. An older mean population age would reduce the mild tension between the models and observations. Constraints such as these on highly evolved, luminous stars are essential for interpreting light from more distant, massive and metal-rich galaxies across the Universe.

The detection of time variation in the integrated light of nearby galaxies opens the way to deriving stellar population ages in these systems by a completely different approach from conventional techniques. In the future, one could imagine high cadence observations of nearby galaxies on >100-d baselines being performed to detect the period distribution of long-period variables by analysis of the power spectra of the time series data. This technique is not limited to old stellar systems; on the contrary, younger systems would show considerably greater temporal variation. For example, on the basis of our models, we expect that 4%, 14%, and 22% of pixels with 10^5 stars would show >1% absolute flux changes over 100 d for ages of 10^{10}, 10^9, and 10^8 yr, respectively. The larger effect at younger ages is due primarily to the larger fractional contribution of long-period variables to the

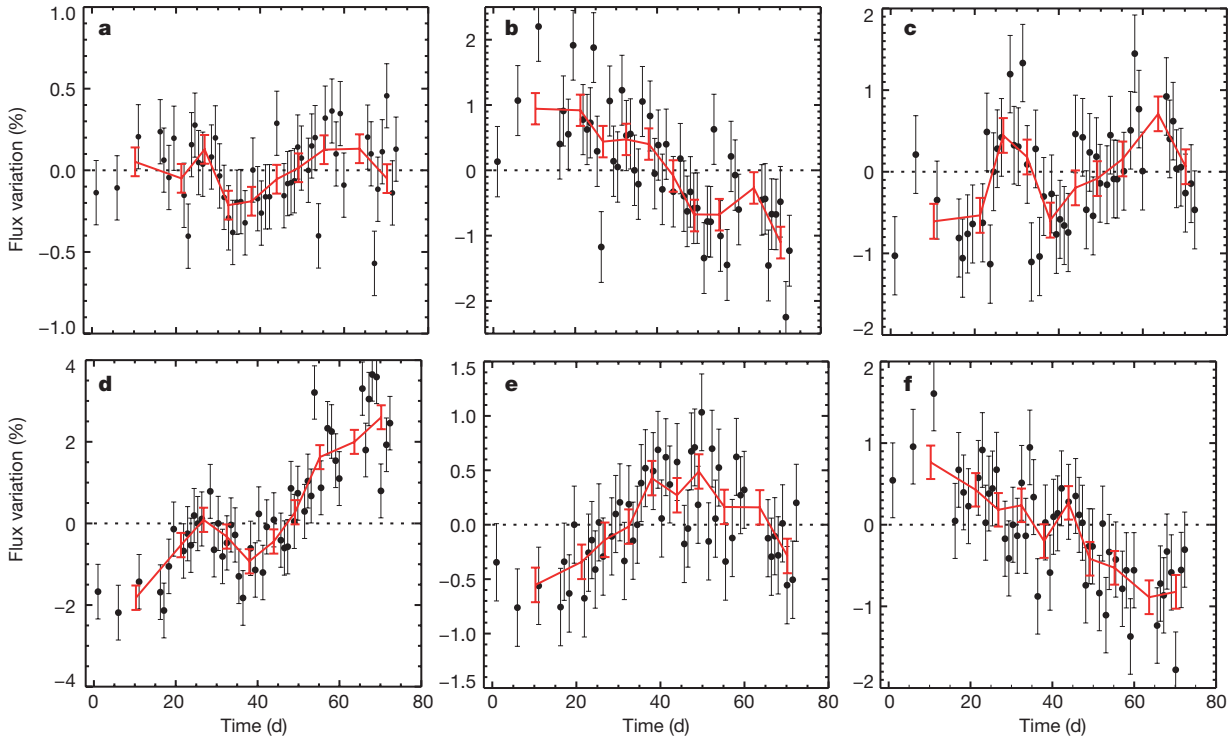


Figure 3 | Observed pixel light curves for M87. a–f, Each panel shows observed relative flux variation over 72 d for a different pixel in M87 (filled black circles). These pixels were selected to highlight the variety of morphology of the light curves, including rising, falling, periodic, and peculiar curves. We unambiguously detect the 'pixel shimmer' due to the contribution of long-period variables to the integrated light. Errors represent 1σ photon counting uncertainties. Red lines and error bars are 5-point boxcar averages of the data.

total light (see Methods for details). It would therefore be relatively straightforward to perform similar studies on nearby spiral galaxies, where the signal would be much stronger. At a basic and fundamental level, each pixel of an observed galaxy varies measurably in time, and this variation encodes unique information on its underlying stellar population.

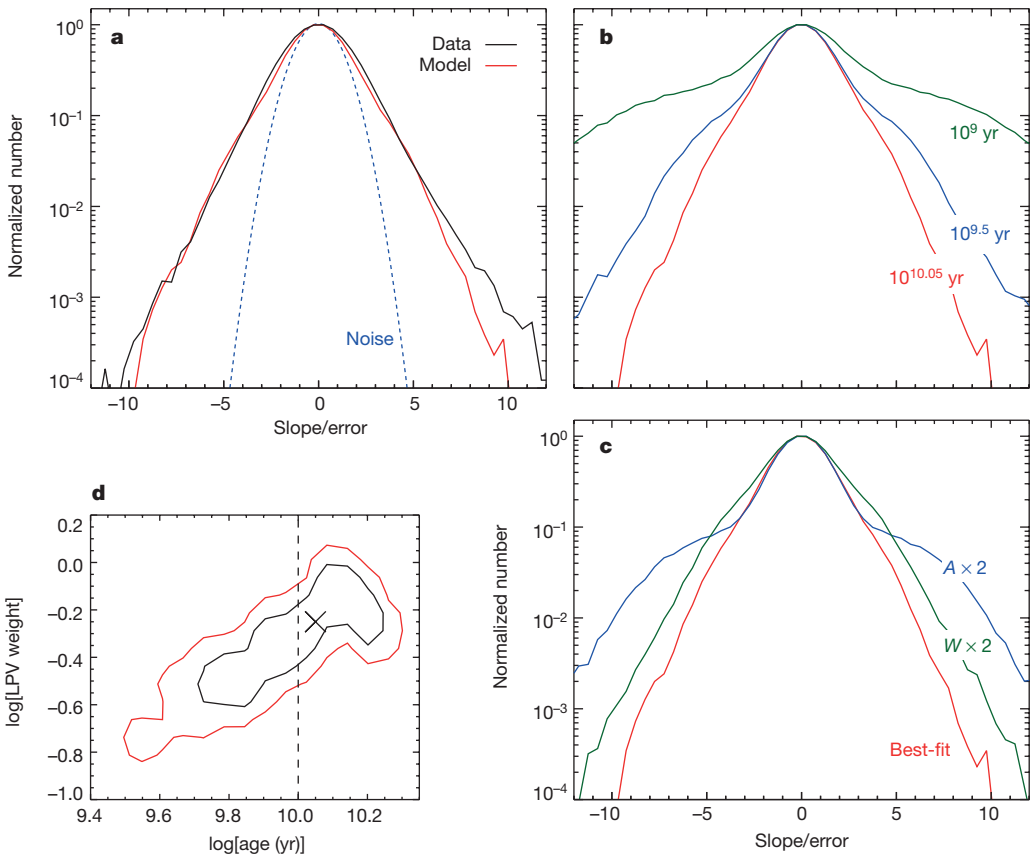


Figure 4 | Statistics of the pixel light curves. a–c, Normalized distributions of the best-fit linear slope of the pixel light curve in units of the 1σ uncertainty on the slope (slope/error). The data (black line in **a**) are compared to several models, including a variable-star-free model (labelled 'noise' in **a**), and models varying the age (**b**), variable star amplitude (**c**), and weight (**c**). Models with varying age and variable star weight ('LPV weight') were fitted to the observed histogram, and the 1σ and 2σ confidence limits on these parameters are shown in **d** (black and red lines, respectively). The best-fit model is shown as a red line in **a–c** and a black cross in **d**. The vertical dashed line in **d** indicates the best-fit age from fitting the integrated light spectrum. In **c** is shown the effect of doubling the amplitudes of the long-period variables ($A \times 2$; blue line) and the weights ($W \times 2$; green line).

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 27 February; accepted 17 September 2015.
Published online 16 November 2015.

1. Fox, M. W. & Wood, P. R. Theoretical growth rates, periods, and pulsation constants for long-period variables. *Astrophys. J.* **259,** 198–212 (1982).
2. Wood, P. R. & Sebo, K. M. On the pulsation mode of Mira variables: evidence from the Large Magellanic Cloud. *Mon. Not. R. Astron. Soc.* **282,** 958–964 (1996).
3. Ita, Y. *et al.* Variable stars in the Magellanic Clouds — II. The data and infrared properties. *Mon. Not. R. Astron. Soc.* **353,** 705–712 (2004).
4. Vassiliadis, E. & Wood, P. R. Evolution of low- and intermediate-mass stars to the end of the asymptotic giant branch with mass loss. *Astrophys. J.* **413,** 641–657 (1993).
5. Tonry, J. & Schneider, D. P. A new technique for measuring extragalactic distances. *Astron. J.* **96,** 807–815 (1988).
6. Tonry, J. L. *et al.* The SBF survey of galaxy distances. IV. SBF magnitudes, colors, and distances. *Astrophys. J.* **546,** 681–693 (2001).
7. Blakeslee, J. P. *et al.* The ACS Fornax cluster survey. V. Measurement and recalibration of surface brightness fluctuations and a precise value of the Fornax-Virgo relative distance. *Astrophys. J.* **694,** 556–572 (2009).
8. van Dokkum, P. G. & Conroy, C. Fluctuation spectroscopy: a new probe of old stellar populations. *Astrophys. J.* **797,** 56 (2014).
9. Groenewegen, M. A. T. & Blommaert, J. A. D. L. Mira variables in the OGLE bulge fields. *Astron. Astrophys.* **443,** 143–156 (2005).
10. Soszyński, I. *et al.* The Optical Gravitational Lensing Experiment. The OGLE-III catalog of variable stars. IV. Long-period variables in the Large Magellanic Cloud. *Acta Astronom.* **59,** 239–253 (2009).
11. Soszyński, I. *et al.* The Optical Gravitational Lensing Experiment. The OGLE-III catalog of variable stars. XV. Long-period variables in the Galactic bulge. *Acta Astronom.* **63,** 21–36 (2013).
12. Salpeter, E. E. The luminosity function and stellar evolution. *Astrophys. J.* **121,** 161–167 (1955).
13. Waters, C. Z., Zepf, S. E., Lauer, T. R. & Baltz, E. A. Color bimodality in M87 globular clusters. *Astrophys. J.* **693,** 463–471 (2009).
14. Peng, E. *et al.* The color-magnitude relation for metal-poor globular clusters in M87: confirmation from deep HST/ACS imaging. *Astrophys. J.* **703,** 42–51 (2009).
15. Bird, S. *et al.* The inner halo of M 87: a first direct view of the red-giant population. *Astron. Astrophys.* **524,** A71 (2010).
16. Conroy, S. & van Dokkum, P. G. The stellar initial mass function in early-type galaxies from absorption line spectroscopy. II. Results. *Astron. Astrophys.* **760,** 71–87 (2012).
17. Gould, A. Search for intracluster MACHOs by pixel lensing of M87. *Astrophys. J.* **455,** 44–49 (1995).
18. Baltz, E. A. *et al.* Microlensing candidates in M87 and the Virgo Cluster with the Hubble Space Telescope. *Astrophys. J.* **610,** 691–706 (2004).
19. Ferrarese, L., Côté, P. & Jordán, A. Hubble Space Telescope observations of novae in M49. *Astrophys. J.* **599,** 1302–1319 (2003).

Acknowledgements We thank M. Groenewegen for discussions. C.C. thanks B. Holden and C. Rockosi for asking the question that provided the spark for this paper: 'can one detect Mira variables in integrated light?'

Author Contributions C.C. constructed the models, led the data processing, and contributed to the analysis and interpretation. P.G.v.D. contributed to the analysis and interpretation. J.C. generated the stellar evolution models and contributed to the analysis and interpretation.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.C. (cconroy@cfa.harvard.edu).

METHODS

Data reduction and tests. Owing to the small expected amplitude of the time-dependent flux signal for M87, great care was taken to control systematic effects. In this section we describe the details of the data reduction procedure and the additional corrections that were applied to the images.

We began with the publicly available HST images, in which the four dithered exposures per visit were combined and astrometrically aligned, resampled by the drizzling process, and cosmic rays were removed. The public images include flat field corrections and a standard sky subtraction. We used five globular clusters to refine the astrometric alignment via subpixel shifts (using bilinear interpolation). The mean shift was 0.25 pixels (in both x and y directions) in the unbinned images. All of our analysis was performed on images binned 4×4, so these shifts are a tiny fraction of the final pixel size.

Owing to the large angular extent of M87, the standard ACS pipeline is not able to accurately measure the true sky background. We therefore applied a correction to the sky subtraction. We assumed that the true M87 surface brightness profile is that reported by Kormendy[20], which was derived by combining a variety of space and ground-based data. Using this profile, we estimated the sky background in our ACS images by minimizing the residuals between the ACS data and Kormendy's profile with the sky background, normalization, and a linear colour gradient as free parameters (the last is to account for differences between our F814W filter and Kormendy's V-band profile). This was done separately for each of the 52 images. We refer to this as the primary sky background correction.

In order to test the fidelity of the images over the 72 d, we selected three background galaxies and measured their fluxes within an 8-pixel aperture. These galaxies should show no detectable temporal variation. The resulting temporal variation of the total flux from these galaxies is shown in Extended Data Fig. 1a. There are no obvious time-dependent trends. However the scatter is 0.5%, which is relatively large compared to the signal of interest (of the order of 1%). We therefore made several additional modifications to the sky background levels in an effort to reduce the scatter.

We identified three additional background galaxies (that is, not the ones used to measure the flux variation in Extended Data Fig. 1) and measured their flux variation over the duration of the observations. Under the assumption that these sources should have no intrinsic flux variation, we determined a sky background correction necessary to bring the flux of the background galaxies to a constant. The average correction determined this way was 0.002 counts s^{-1}. At this point in the analysis, the distribution of pixel light curve slopes showed a slight preference for positive slopes (the mean slope/error was +0.5). Under the assumption that the true distribution should have a mean of zero, we subtracted a linearly varying sky background component (which scaled as $5 \times 10^{-5} t$). These two corrections yield a distribution of pixel light curve slopes with zero mean (by construction), and a temporal flux variation in the three reference background galaxies with a scatter of 0.2% as shown in Extended Data Fig. 1b. Moreover, a 5-point boxcar average of the light curve of the background galaxies shows flux variation at the $\lesssim$0.1% level. From this test we conclude that it should be possible to measure intrinsic flux variation at the sub-percent level, at least for pixels where photon counting noise is not the dominant source of uncertainty.

We emphasize that the additional sky background corrections discussed above do not materially change our conclusions. While these corrections result in a shift in the histogram of slope/error values, they have no effect on the width of the distribution. Moreover, the example pixel light curves shown in Fig. 2 are unchanged within their 1σ error bars.

Approximately half of the exposures were obtained at a detector location offset by 60–70 pixels compared to the other half of the exposures. This provides a further test that the trends shown in Fig. 3 and the statistics in Fig. 4 are not dominated by unknown systematics at the level of the detector; if they were, one would have expected to see flux variation that correlated with the dither pattern, but such correlations, if present, are within the noise limits of the data.

As a final test of both the data reduction and our results, in Extended Data Fig. 2 we show histograms of the flux variation over the 72-d observing window. The flux variation was computed by temporally binning the exposures by five to reduce the Poisson noise in the measurement and computing (maximum−minimum)/mean flux at each pixel. The results are shown for three bins of pixel fluxes (the legends show the cuts in units of counts per second and the total number of pixels per bin). The data (black lines) are compared to our best-fit model as derived from fitting the pixel light curve slopes (red lines) and a model without long-period variables (blue lines). The good agreement between the model and data in all three panels is a strong indication that our measurements are reliable, as the panels probe a factor of 40 in dynamic range in pixel fluxes. Systematic issues with, for example, the sky subtraction would show up most strongly in the pixels with low count rates, and yet the observations and models agree very well in that regime. Moreover, this flux variation metric is model independent and so the difference between the variable-star-free model and the observations provides further strong support that the variation detected in the observations is real and not an artefact of some unknown systematics. There do exist subtle differences between the model and data that vary as a function of the pixel flux, but this could be due to changes in the underlying stellar populations as the pixels with low fluxes are in the outskirts, where the ages and metallicities of the stars are expected to differ from the central regions.

Modelling long-period variables. Here we provide additional details regarding the incorporation of long-period variables in the stellar population synthesis modelling. We start with stellar isochrones that include all relevant evolutionary phases, including thermally pulsating asymptotic giant branch (AGB) stars. We include a model for circumstellar dust around these stars, which results in dimmer stars especially for the most intrinsically luminous and evolved stars[21]. Periods (in days) are assigned according to the following equation:

$$\log P = -2.07\log(R/R_{\odot}) - 0.9\log(M/M_{\odot}) \quad (1)$$

which assumes that the stars pulsate in their fundamental mode[4]. Next, we require a relation between pulsation period and amplitude. This relation is shown in Extended Data Fig. 3 for stars in the Galactic bulge from OGLE data[11]. Symbols are colour-coded according to the type of pulsator. The dashed lines are the adopted period–amplitude (P–A) sequences:

$$\log A = 0.5\log P - 1.25 \quad (2)$$

for the Mira sequence ($\log P > 2.2$), and:

$$\log A = 2\log P - 5 \quad (3)$$

for the semi-regular variable (SRV) sequence ($1.0 < \log P < 2.2$). In the equations above, P is in days and the amplitude A is in the I band in magnitudes. We note that the SRV sequence is included for completeness but has a very small effect on the model predictions.

The equations above, along with the initial mass function weights determined by the masses of the AGB stars in the isochrones, completely specify our default variable-star model. In order to convert fluxes to luminosities, we have assumed a distance to M87 of 16.7 Mpc (ref. 7). In order to explore the constraining power of the data, we considered variation in both the amplitude of the long-period variables, implemented as an overall scaling of all the amplitudes by the same factor, and the weight given to the variable stars in the population synthesis. The latter can be interpreted as a change to the typical lifetime.

We have taken great care to ensure that the long-period variable phase is well-resolved in the isochrone tables. The isochrones were constructed from 185 individual mass models and with 600 equivalent evolutionary points in the thermally pulsating AGB phase alone. We have run a variety of tests to ensure that our model predictions are 'converged'; for example, we have created models with fewer evolutionary points and fewer input mass models and the resulting predictions are very similar. For context, at 10 Gyr our isochrones contain 350 points on the AGB with periods >200 d, while the publicly available Padova[22] isochrones contain only 3 such points.

Extended Data Figure 4 quantifies the fractional contribution of long-period variables to the total flux of a stellar population as a function of wavelength, age, and metallicity ([Z/H]). The flux contribution peaks in the age range of $10^{8.5}$–10^{9} yr and increases towards redder bands. The trend with wavelength is a reflection of the fact that variable stars are cool and so emit most of their light in the near-infrared. We caution that the wavelength-dependence shown here does not directly translate into the wavelength-dependence of the time-dependent signal because the period–amplitude relation also depends on wavelength. As the pulsation directly affects the radius and hence the temperature, for these cool stars one expects and indeed observes that the amplitudes are larger in the bluer wavebands[23]. The metallicity-dependence is relatively modest, at least over the range [Z/H] = −0.3 to [Z/H] = +0.3, typical of massive galaxies. It is difficult to provide a simple explanation of the model metallicity variation, as it depends not only on the variable-star lifetime, luminosity, and temperature, but also on the properties of the underlying stellar population.

We do not expect metallicity to play a critical role in the interpretation of the observations for several reasons. First, as noted in the previous paragraph, the models suggest a relatively weak metallicity-dependence of the long-period variable flux contribution. Second, M87 harbours a metallicity gradient[24], extending from slightly super-solar in the inner $R_e/8$ to slightly sub-solar at R_e, where R_e is the effective radius. Despite this metallicity gradient, our best-fit model provides an equally good fit to the pixel shimmer statistics in both the central region and the outskirts, as shown in Extended Data Fig. 2.

We note here that individual long-period variables have been detected in nearby galaxies including the Magellanic Clouds[25], M31[26], and M32[27]. The most distant galaxy with secure detections of individual long-period variable stars is NGC 5128[28], and in this case the observations were confined to the outskirts where the stellar density was sufficiently low to permit the separation of the brightest evolved stars from the background sea of lower luminosity stars. These observations of individual long-period variables should provide very useful constraints on the modelling of such stars, and we intend to make use of these constraints in future work.

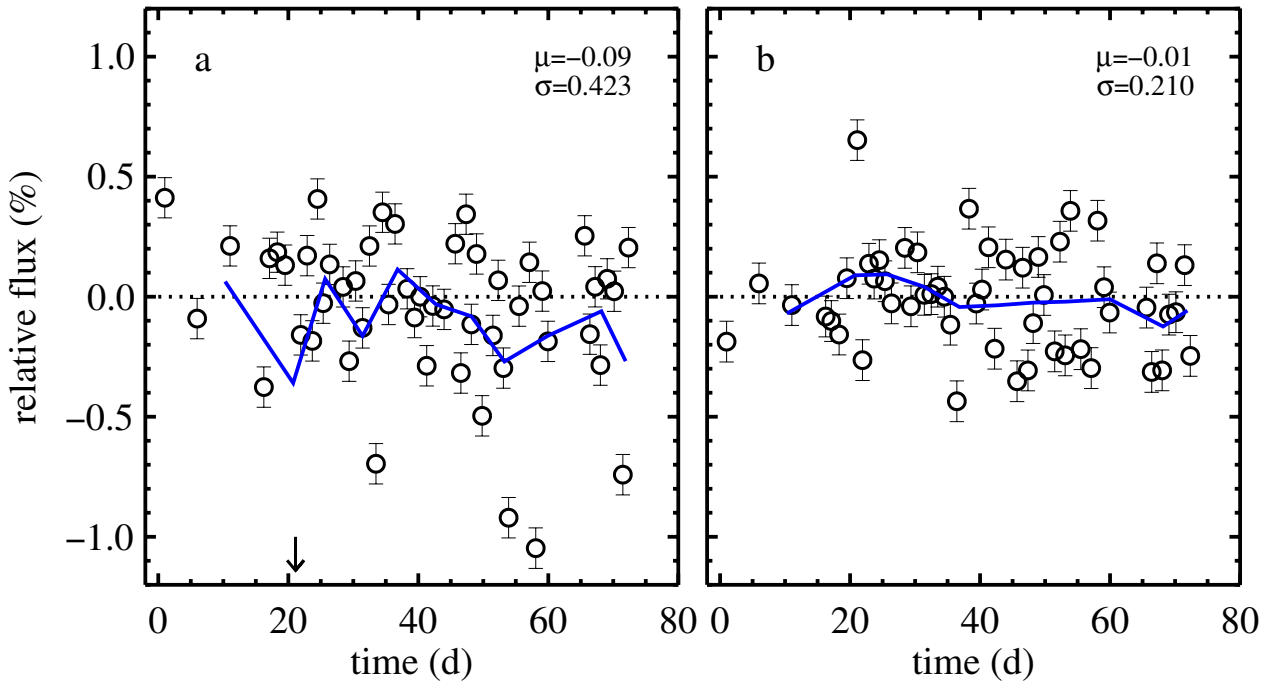
Trends with radius. The HST field of view covers the central $3.3' \times 3.3'$ of M87, of which the inner $\sim 1' \times 1'$ has a signal-to-noise ratio $S/N \gtrsim 100$ per pixel for the observations that were analysed herein. Kormendy[20] reports an effective radius of $R_e = 3.2'$ so the region of the images with high S/N covers the inner $\sim 0.3 R_e$. Extended Data Figure 5 shows several important quantities as a function of R/R_e for our best-fit model of M87. Extended Data Figure 5a shows the stellar mass per pixel for the underlying smooth stellar distribution. Extended Data Figure 5b shows the fraction of pixels with $|\text{slope/error}| > 2$. In the main text we reported that 24% of pixels reach this criterion, and in fact that percentage remains approximately constant with radius. The constancy is the result of two opposing effects: at larger radius the number of stars per pixel is lower, which implies a larger variable-star signal. The effect on the slope scales approximately as $\sqrt{N}$ (here N is the number of stars per pixel) as multiple variable stars with random phases will cancel each other out in a central-limit-theorem-like process. However, at larger radius the S/N is lower, and for a fixed exposure time this also scales as $\sqrt{N}$. Thus, for a fixed exposure time, the detectability of long-period variables is fairly constant with radius.

Extended Data Figure 5c and d show the model trends with radius for a noise-free model (infinite S/N). In this case it is clear that the absolute effect of long-period variables is larger at larger radius. Extended Data Figure 5c shows the fraction of pixels with >1% peak-to-peak flux variation over 200 d. Old stellar populations with a pixel mass $< 10^6 M_\odot$ yield >1% flux variation in ~10% of the pixels. Extended Data Figure 5d compares the surface brightness fluctuation (SBF) amplitude at a single epoch to the mean temporal variation over a 200-d baseline; the latter is smaller than the former by a factor of ~5. The SBF amplitude is computed as the standard deviation of the model flux divided by a smooth model for the flux.

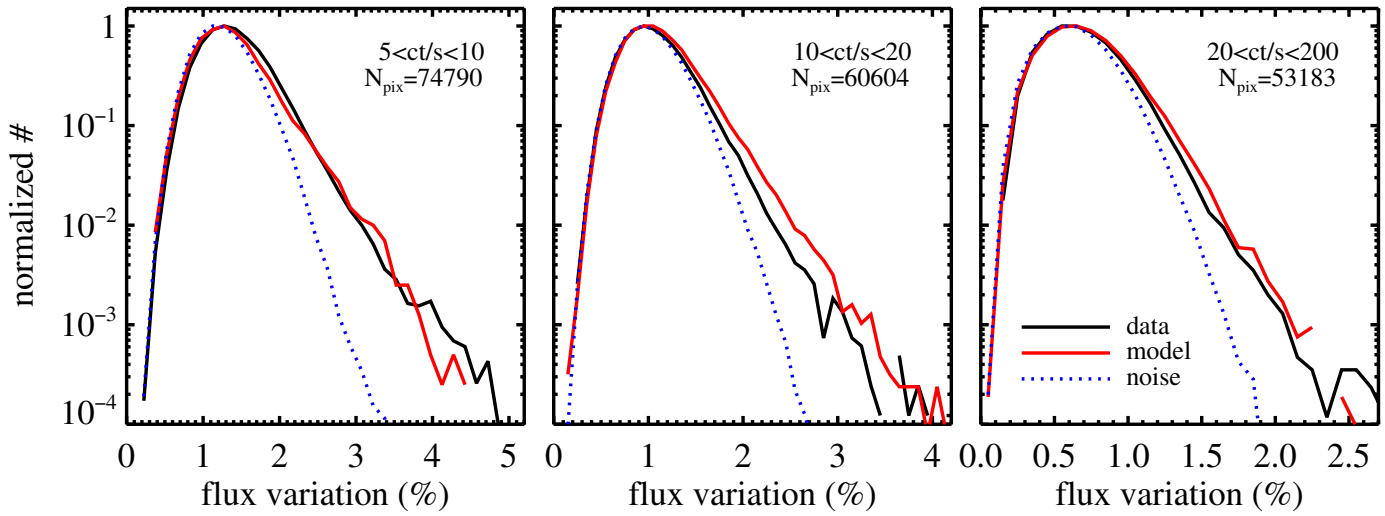
We close by noting that while the overall effect of long-period variables on the integrated light is relatively modest at old stellar ages, it is much more prominent for younger stellar populations, for example, in the 10^8–10^9 yr range. Future work devoted to younger stellar populations will therefore probably uncover a rich array of observational signatures of time variable stellar populations.

Code availability. We have opted not to make the code used in this manuscript available because the data reduction and analysis is fairly straightforward and can be easily reproduced following the methods described herein.

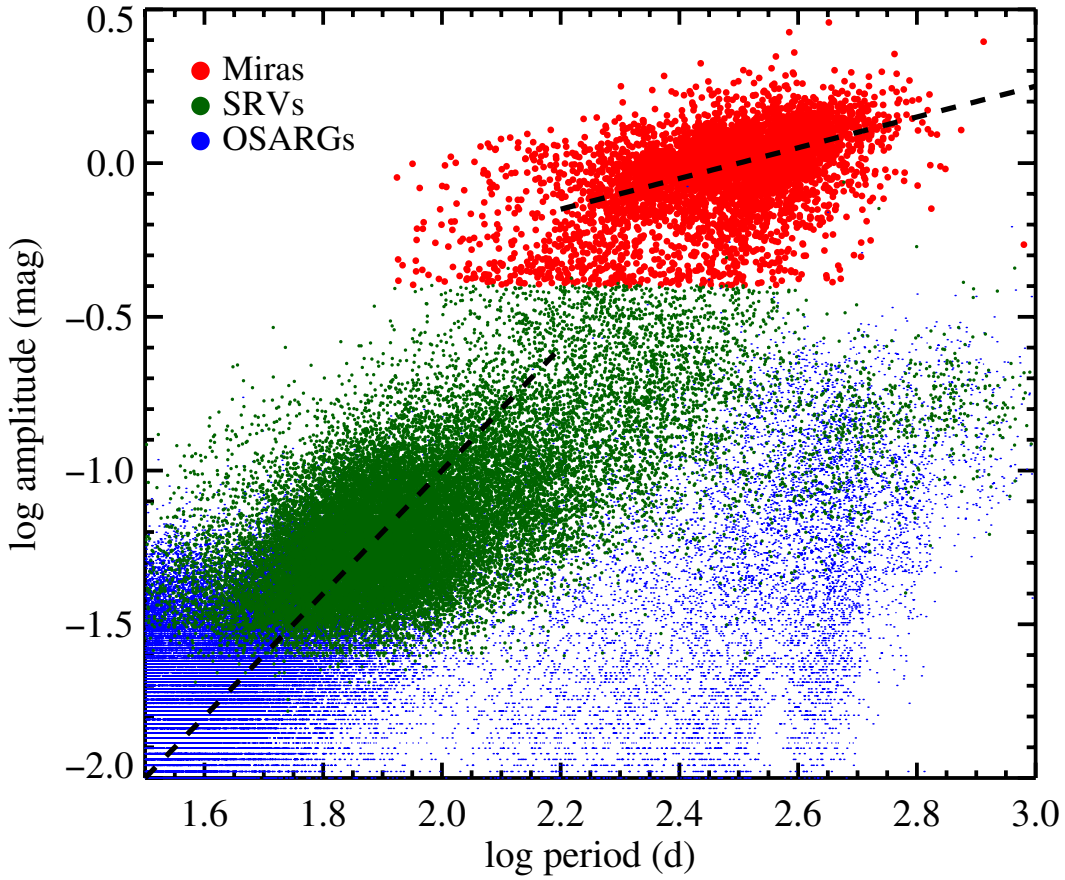
20. Kormendy, J., Fisher, D. B., Cornell, M. E. & Bender, R. Structure and formation of elliptical and spheroidal galaxies. *Astrophys. J.* **182** (Supp.), 216–309 (2009).
21. Villaume, A., Conroy, C. & Johnson, B. Circumstellar dust around AGB stars and implications for infrared emission from galaxies. *Astrophys. J.* **806,** 82 (2015).
22. Marigo, P. *et al.* Evolution of asymptotic giant branch stars. II. Optical to far-infrared isochrones with improved TP-AGB models. *Astron. Astrophys.* **482,** 883–905 (2008).
23. Smith, B. J., Leisawitz, D., Castelaz, M. W. & Luttermoser, D. Infrared light curves of Mira variable stars from COBE DIRBE data. *Astron. J.* **123,** 948–964 (2002).
24. Kuntschner, H. *et al.* The SAURON project — XVII. Stellar population analysis of the absorption line strength maps of 48 early-type galaxies. *Mon. Not. R. Astron. Soc.* **408,** 97–132 (2010).
25. Wood, P. R., Bessell, M. S. & Fox, M. W. Long-period variables in the Magellanic Clouds — supergiants, AGB stars, supernova precursors, planetary nebula precursors, and enrichment of the interstellar medium. *Astrophys. J.* **272,** 99–115 (1983).
26. Fliri, J., Riffeser, A., Seitz, S. & Bender, R. The Wendelstein Calar Alto Pixel-lensing Project (WeCAPP): the M 31 variable star catalogue. *Astron. Astrophys.* **445,** 423–439 (2006).
27. Davidge, T. J. & Rigaut, F. Photometric variability among the brightest asymptotic giant branch stars near the center of M32. *Astrophys. J.* **607,** L25–L28 (2004).
28. Rejkuba, M., Minniti, D. & Silva, D. R. Long period variables in NGC 5128. I. Catalogue 2. *Astron. Astrophys.* **406,** 75–85 (2003).



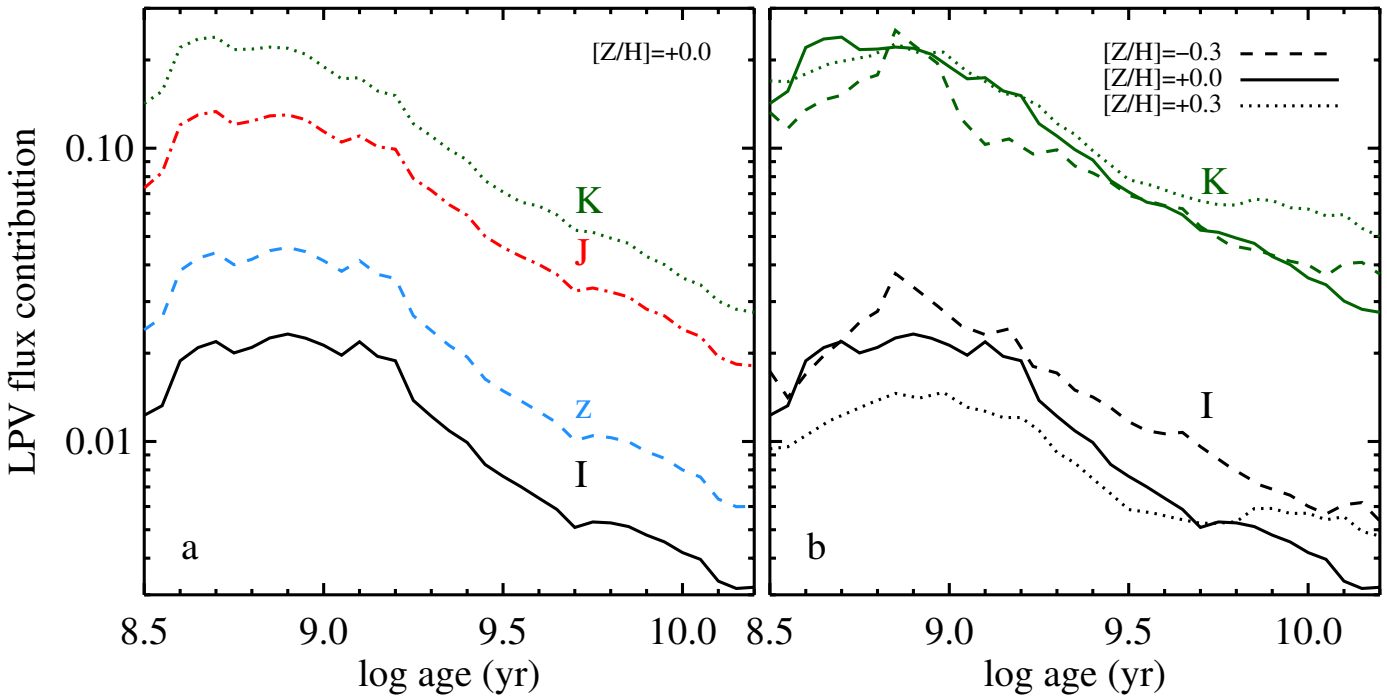
Extended Data Figure 1 | Flux of background galaxies. Shown is the time variation of the flux of three background galaxies. The background galaxies should show no intrinsic time variation in their flux and therefore serve as a test of the stability of the data. The mean (μ) and standard deviation (σ) are reported in each panel. The 1σ error on each point due to photon counting uncertainty is 0.09%. The solid line is a 5-point boxcar average. **a**, Flux variation after the standard data reduction including the primary sky background correction. The arrow indicates a point that lies at −2.1. **b**, Flux variation after additional corrections were applied to the sky background levels. These additional corrections allow us to achieve a stability of ~0.1% for boxcar-averaged time series data.



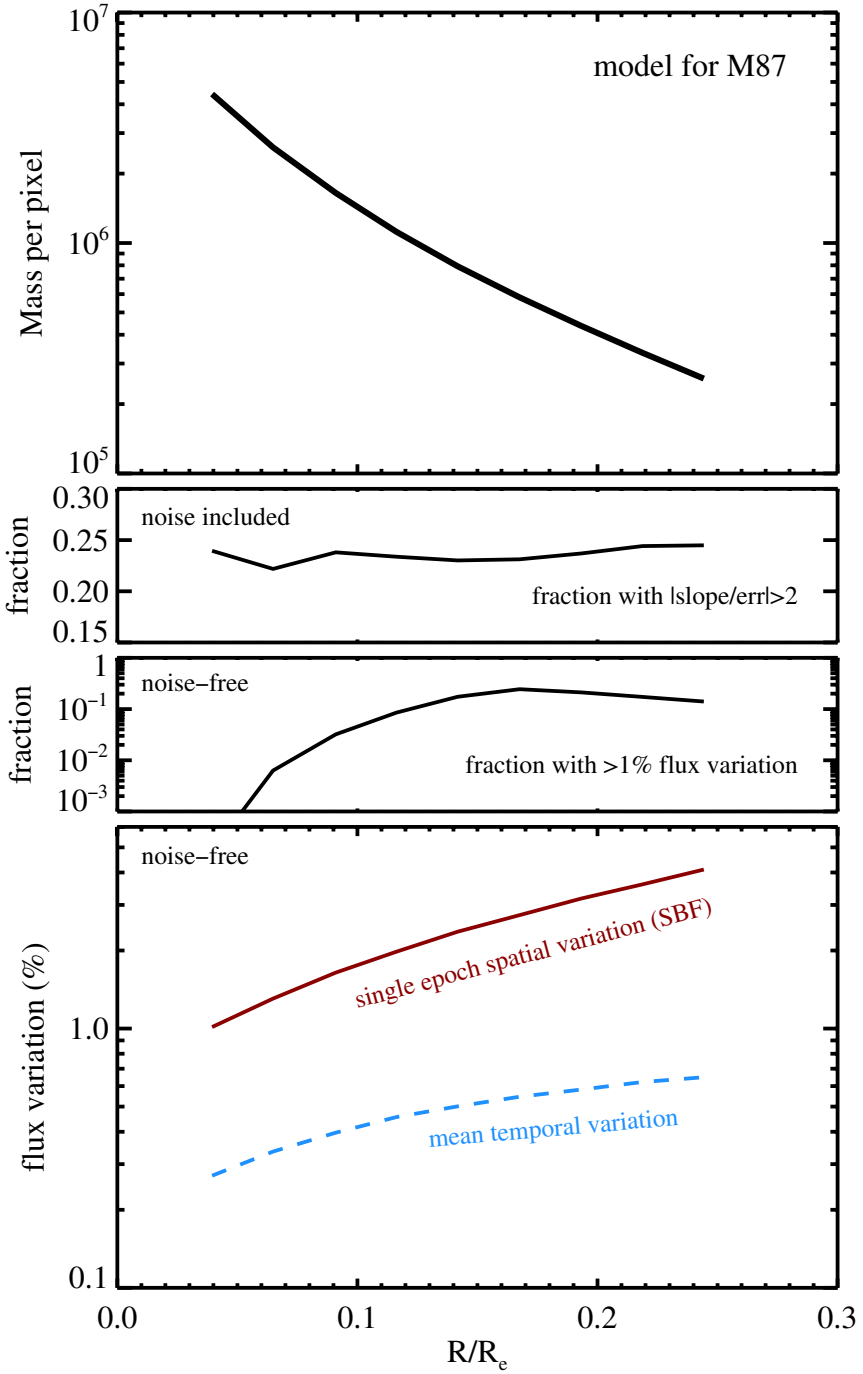
Extended Data Figure 2 | Flux variation distributions. Shown is the normalized distribution of (maximum−minimum)/mean fluxes over the 72 observing windows, separated into three bins of counts per second (ct/s). The data (black lines) are compared to the best-fit model (red lines) and a noise-only model (blue dotted lines). Also shown in each panel is the number of pixels, N_{pix}, contributing to the distribution.



Extended Data Figure 3 | Amplitude versus period for luminous variable stars. Data are for Galactic bulge stars from the OGLE survey[11] measured in the I band. The distinct classes of Miras, semi-regular variables (SRVs), and OGLE small-amplitude red giants (OSARGs) are shown as red, green, and blue symbols. Lines are the adopted sequences for Miras and SRVs; these relations are used to assign pulsation amplitudes in our model.



Extended Data Figure 4 | Long-period variable (LPV) star flux contribution versus age, wavelength, and metallicity. a, Fractional contribution to the total luminosity versus age in four bandpasses: I (0.8 μm), z (0.9 μm), J (1.2 μm), and K (2.4 μm). The flux contribution scales approximately as $t^{-1/2}$. **b**, Flux contribution versus age and metallicity for the I and K bandpasses. The metallicity range shown encompasses the observed variation in M87 within R_e.



Extended Data Figure 5 | Radial variation of model properties for M87. **a**, Stellar mass per pixel for the smooth underlying model for M87 as a function of R/R_e where R_e is the effective radius. **b**, Fraction of pixels with |slope/error| > 2. **c**, Fraction of pixels in a noise-free model with >1% peak-to-peak flux variation over 200 d. **d**, Strength of surface brightness fluctuation (SBF) signal at a single epoch compared to the mean temporal variation due to variable stars in a noise-free model.

LETTER

doi:10.1038/nature16137

Collisionless encounters and the origin of the lunar inclination

Kaveh Pahlevan[1] & Alessandro Morbidelli[1]

The Moon is generally thought to have formed from the debris ejected by the impact of a planet-sized object with the proto-Earth towards the end of planetary accretion[1,2]. Models of the impact process predict that the lunar material was disaggregated into a circumplanetary disk and that lunar accretion subsequently placed the Moon in a near-equatorial orbit[3–6]. Forward integration of the lunar orbit from this initial state predicts a modern inclination at least an order of magnitude smaller than the lunar value—a long-standing discrepancy known as the lunar inclination problem[7–9]. Here we show that the modern lunar orbit provides a sensitive record of gravitational interactions with Earth-crossing planetesimals that were not yet accreted at the time of the Moon-forming event. The currently observed lunar orbit can naturally be reproduced via interaction with a small quantity of mass (corresponding to 0.0075–0.015 Earth masses eventually accreted to the Earth) carried by a few bodies, consistent with the constraints and models of late accretion[10,11]. Although the encounter process has a stochastic element, the observed value of the lunar inclination is among the most likely outcomes for a wide range of parameters. The excitation of the lunar orbit is most readily reproduced via collisionless encounters of planetesimals with the Earth–Moon system with strong dissipation of tidal energy on the early Earth. This mechanism obviates the need for previously proposed (but idealized) excitation mechanisms[12,13], places the Moon-forming event in the context of the formation of Earth, and constrains the pristineness of the dynamical state of the Earth–Moon system.

The Moon-forming impact is thought to have generated a compact circumplanetary disk (within ten Earth radii, R_E) of debris out of which the Moon rapidly accreted. Like Saturn's rings, the proto-lunar disk would be expected to become equatorial on a timescale that is rapid relative to its evolutionary timescale. Hence, so long as the proto-lunar material disaggregated into a disk following the giant impact, the Moon is expected to have accreted within about one degree of the Earth's equatorial plane[6]. Tidal evolution calculations suggest that for every degree of inclination of the lunar orbital plane relative to the Earth's equatorial plane at an Earth–Moon separation of $10R_E$, the current lunar orbit would exhibit about half a degree of inclination relative to Earth's orbital plane[7–9]. The modern lunar inclination of approximately 5° would—without external influences—translate to an inclination of about 10° to Earth's equatorial plane at $10R_E$ shortly after lunar accretion. This approximately tenfold difference between theoretical expectations of lunar accretion and the observed Earth–Moon system is known as the lunar inclination problem.

Previous work on this problem has sought to identify mechanisms such as a gravitational resonance between the newly formed Moon and the Sun[12] or the remnant proto-lunar disk[13] that can excite the lunar inclination to a level consistent with its current value. Neither of these scenarios is satisfactory, however, as the former requires particular values of the tidal dissipation parameters and the latter has only been shown to be viable in an idealized system in which a single, fully formed Moon interacts with a single pair of resonances in the proto-lunar disk. Moreover, previous works assumed that the excitation of the lunar orbit was determined during interactions that essentially coincided with lunar origin. Here, we propose that the lunar inclination arose much later as a consequence of the sweep-up of remnant planetesimals in the inner Solar System.

After the giant impact and at most 10^3 years[14,15], the Moon has accreted, interacted with[13] and caused the collapse of the remnant proto-lunar disk onto the Earth[6], passed the evection resonance with the Sun[3,12,16], and begun a steady outward tidal evolution. On a timescale (10^6–10^7 years) that is rapid relative to that characterizing depletion of planetesimals in the final post-Moon-formation stage of planetary accretion[17] (called 'late accretion'), the lunar orbit expands, owing to the action of tides, to an Earth–Moon separation of $20R_E$–$40R_E$. During this time, the lunar orbit transitions from precession around the spin axis of Earth to precession around the normal vector of the heliocentric orbit[8], and its inclination becomes insensitive to the shifting of the Earth's equatorial plane via subsequent accretion[18]. However, as we show, lunar inclination becomes more sensitive to gravitational interactions with passing planetesimals as the tidal evolution of the system proceeds. The sensitivity is such that it renders the lunar orbital excitation a natural outcome of the sweep-up of the leftovers of accretion and yields a new constraint on the dynamical and tidal environment of the Earth–Moon system in the 10^8 years immediately following its origin.

Although subsequent collisions with the Earth–Moon system have been previously considered as a mechanism for dynamical excitation[18], the collision of inner Solar System bodies with the Earth tends to be preceded by a large number (10^3–10^4) of collisionless encounters. Excitation via this process is governed by two relevant timescales: the timescale over which remnant populations in the inner Solar System are lost via accretion onto the planets and the Sun (several tens of millions of years[17]); and the timescale for the lunar tidal orbital expansion, which is a rapidly varying function of the Earth–Moon distance. The Earth–Moon distance is important because it determines the system cross-section for collisionless encounters with remnant bodies. Accordingly, the rate of tidal expansion of the lunar orbit during the first approximately 10^8 years after the giant impact is also important. As tidal evolution proceeds and the Earth–Moon separation increases, the system becomes increasingly susceptible to collisionless excitation, while populations capable of exciting the system are progressively depleted. A few tens of millions of years after the Moon-forming event, the Earth–Moon system reaches an optimal capacity for excitation via gravitational encounters: a dynamically excitable system co-existing with a substantial remnant-body population.

Here, we run a series of Monte Carlo simulations to set constraints on the outcome of repeated encounters of massive bodies with the evolving Earth–Moon system. The simulations are carried out until the populations are exhausted either through collision with the Earth or through non-terrestrial loss channels (for details, see Methods). A sample run of dynamical excitation during the first approximately 10^8 years of Earth–Moon history is shown in Fig. 1. No single event dominates: several strong encounters contribute substantially to the

[1]Laboratoire Lagrange, Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS Boulevard de l'Observatoire CS 34229, 06304 Nice Cedex 4, France.

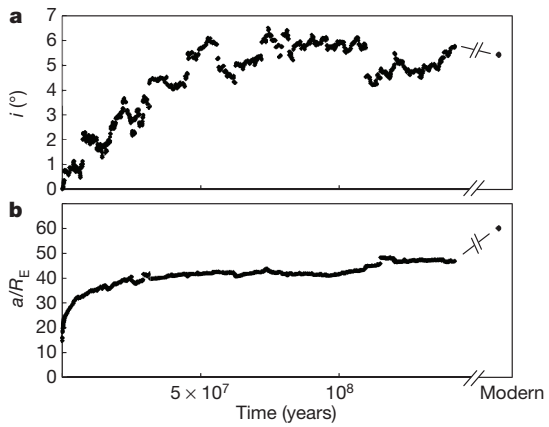


Figure 1 | Sample realization. a, **b**, A model of the early lunar orbit subject to tidal evolution ($k_2/Q = 0.1$) and encounters leading to collision of two $0.00375M_E$ bodies with the Earth. The semi-major axis of the evolving lunar orbit a is given in Earth radii (shown in **b**). Although not every encounter increases the lunar inclination i, the cumulative effect shows a tendency towards excitation (shown in **a**). Notable interactions include merging collisions with the Earth kicking the lunar orbit via recoil (at 29.1 Myr and 31.5 Myr), several exceptionally close encounters with the Moon (at 7.3 Myr and 109.6 Myr) and the exhaustion of the population (at 141.7 Myr) ultimately marking the end of the simulation. Subsequent inclination damping owing to planetary tides is modest (from 5.8° at $47R_E$ at the end of the simulation to 5.4° at $60R_E$), a feature that is typical of this 'late' excitation mechanism.

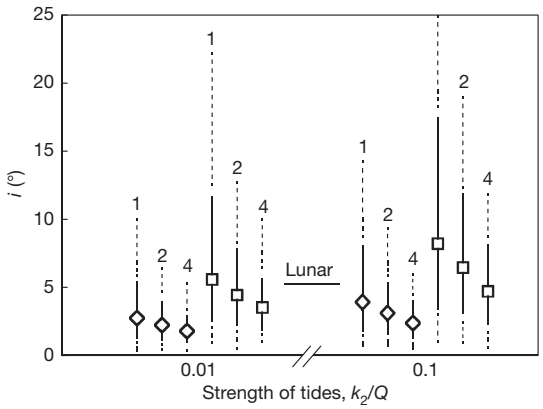


Figure 2 | Summary of simulations. Median values (symbols), and 1σ (solid lines) and 2σ (dashed lines) intervals for the lunar inclination i at the end of the simulations after damping via planetary tides to the modern Earth–Moon separation. The excitation in the modern lunar orbit is plotted for comparison, indicated by the horizontal line labelled 'lunar'. Diamonds correspond to simulations with $0.0075M_E$ accreted to Earth; squares correspond to $0.015M_E$. 'Strong' tidal dissipation ($k_2/Q = 0.1$) corresponds to a hot dissipative silicate Earth, and 'weak' dissipation ($k_2/Q = 0.01$) represents the geologic average value dominated by dissipation in shallow oceans. In these simulations, the accretion of $0.0075M_E$ (with 'strong' tides) to $0.015M_E$ (with 'weak' tides) frequently reproduces the excitation in the lunar orbit. The number of bodies delivering the late accreted mass in each set of simulations is reported above each symbol.

final excitation. The size distribution of the late-accreting population is assumed to be top heavy, with most of the mass contained in a few massive bodies, as has been previously proposed to explain terrestrial late accretion[10,11]. This particular simulation ultimately results in two $0.00375M_E$ planetesimals (where M_E is Earth's mass) left over from the formation process colliding with the Earth. Tidal damping of the lunar inclination is applied along with the lunar orbital expansion, following equation (1) (see Methods). We do not consider the possibility that the lunar inclination might have been more strongly damped via dissipation in the lunar magma ocean, as recently proposed[19]. In the Methods, we show that this effect is not important as long as the lunar magma ocean crystallized within a few tens of millions of years.

Lunar orbital excitation in this epoch depends on the total mass of leftover planetesimals, the number of bodies carrying the mass and their orbital distribution, the rate of terrestrial tidal dissipation, and a stochastic element. In Fig. 2, we show results of simulations of the excitation of the lunar inclination due to interaction of the system with a small amount of mass (equivalent to $0.0075M_E$–$0.015M_E$ eventually accreted to the Earth), for different values of the strength of tidal dissipation and the number of bodies delivering the mass (which is constrained to be <20 colliding with Earth via models of late accretion[10,11]).

Several features are apparent. First, there is a quasi-linear dependence of the excitation on the total mass of late accretion: other variables being equal, excitation corresponding to $0.015M_E$ of late accretion is approximately twice as great as that with $0.0075M_E$. The mass accreted onto the Earth thereby provides a proxy for collisionless excitation. Second, the lunar orbital excitation exhibits some dependence on the strength of tidal evolution: stronger dissipation within the Earth drives the lunar orbit outwards faster and exposes the system to more collisionless events. Simulations with the weakest tides that we considered—characterized by the ratio of the tidal Love number to the specific dissipation function $k_2/Q = 0.01$—typically excite the lunar inclination with a planetesimal population consistent with $0.015M_E$ of late accretion, whereas with stronger tidal dissipation ($k_2/Q = 0.1$), lunar inclination is routinely excited by a planetesimal population carrying $0.0075M_E$ of late accretion. Third, there exists a negative dependence of the excitation on the number of bodies involved in late accretion, such that the mechanism requires a population that is top heavy (with most of the mass delivered via the most massive bodies). For a given mass of late accretion, a greater number of bodies also renders the distribution of lunar inclinations more strongly peaked and predictions of the expected excitation more precise. Despite an order of magnitude of uncertainty in the strength of early terrestrial tides (k_2/Q) and in the number of bodies involved in the leftover population, and the stochasticity that is inherent in collisionless encounters, close encounters with a population of planetesimals delivering $0.0075M_E$–$0.015M_E$ to the Earth after the Moon-forming event can robustly reproduce the excitation that characterizes the lunar orbit.

The angular momentum of the Earth–Moon system at the time of its origin is a central feature diagnostic of various proposed giant-impact scenarios[1,3,4,18]. Given that the lunar orbit provides a sensitive dynamical measure of encounters with the Earth after the origin of the Moon, we ask whether such gravitational interactions were effective in injecting or extracting angular momentum. Figure 3 summarizes the angular momentum change versus the final excitation. The change in angular momentum corresponding to the modern inclination excitation of approximately 5° is probably a few tens per cent or less. Hence, the standard giant-impact scenario[4] followed by little subsequent dynamical modification is compatible with the dynamical state of the modern system, whereas a high-angular-momentum impact scenario[3,5] would require another dynamical mechanism such as the evection resonance[3,12,16] to be reconciled with the modern Earth–Moon system.

The sensitivity of the orbits of impact-generated satellites to ongoing accretion onto the host planet has several consequences. The degree of orbital excitation resulting from interaction with, and accretion of, $0.0075M_E$–$0.015M_E$ onto the early Earth suggests that collisionless encounters with massive bodies—such as the Moon-to-Mars mass embryos thought to have played a key role in the accretion of the Earth—would have excited satellites to very eccentric orbits ultimately leading to their dynamical loss, either via collision with the host planet or liberation into heliocentric orbit. Such excitability of impact-generated satellite orbits may explain several features of the inner Solar System that have yet to be understood. For example, despite impact-generated satellites being a quasi-generic feature of terrestrial planet formation via giant impact, the absence of an impact-generated

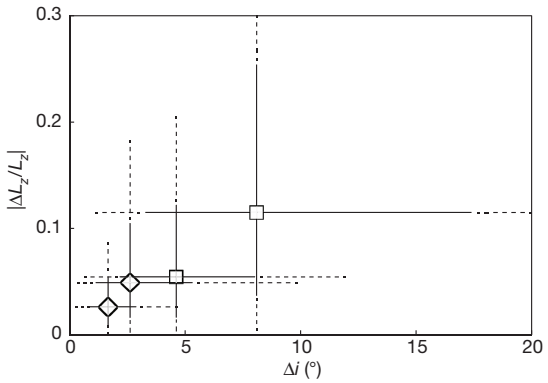


Figure 3 | Angular momentum change of the Earth–Moon system. Median values (symbols), and 1σ (solid lines) and 2σ (dashed lines) intervals for inclination i and angular momentum L_z change via post-lunar collisionless encounters. Diamonds represent realizations with weak tides ($k_2/Q = 0.01$) and $0.0075M_E$ accretion; squares correspond to realizations with stronger dissipation ($k_2/Q = 0.1$) and $0.015M_E$ accretion, bracketing the range in our simulations. Each suite of simulations is composed of two subsets: one with late accretion delivered via one body (greater excitation); the other, four bodies (lesser excitation). Intermediate outcomes with two accreted bodies are omitted for clarity. For a level of excitation consistent with the modern lunar orbit (5.15°), the amount of system angular momentum change is probably $<20\%$. However, the 2σ intervals for the strongest excitation case plotted extend to $\Delta i = 42°$ and $|\Delta L_z/L_z| = 0.48$, implying a small probability ($<5\%$) for angular momentum change $>50\%$.

satellite around Venus[20] and the apparent absence of a pre-Moon terrestrial satellite[21] can be understood: any such early-formed satellites would have been lost via encounters with extant planetary embryos, including perhaps the Moon-forming impactor itself. Moreover, the occurrence of the Moon-forming giant impact late in the history of Earth accretion can be understood as a necessity for its survival: even satellites generated moderately earlier would have been readily dynamically destabilized. Just as the survival of planets depends on the surrounding stellar environment[22], the survival of an impact-generated satellite depends on the planetary environment at the time of origin.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 23 March; accepted 6 October 2015.

1. Cameron, A. G. W. & Ward, W. R. The origin of the Moon. *Lunar Planet. Sci. Conf.* **7,** abstr. 120–122 (1976).
2. Hartmann, W. K. & Davis, D. R. Satellite-sized planetesimals and lunar origin. *Icarus* **24,** 504–515 (1975).
3. Cuk, M. & Stewart, S. T. Making the Moon from a fast-spinning Earth: a giant impact followed by resonant despinning. *Science* **338,** 1047–1052 (2012).
4. Canup, R. M. & Asphaug, E. Origin of the Moon in a giant impact near the end of the Earth's formation. *Nature* **412,** 708–712 (2001).
5. Canup, R. M. Forming a Moon with an Earth-like composition via a giant impact. *Science* **338,** 1052–1055 (2012).
6. Ida, S., Canup, R. M. & Stewart, G. R. Lunar accretion from an impact-generated disk. *Nature* **389,** 353–357 (1997).
7. Mignard, F. The lunar orbit revisited, III. *Moon Planets* **24,** 189–207 (1981).
8. Goldreich, P. History of lunar orbit. *Rev. Geophys.* **4,** 411–439 (1966).
9. Touma, J. & Wisdom, J. Evolution of the Earth–Moon system. *Astron. J.* **108,** 1943–1961 (1994).
10. Bottke, W. F., Walker, R. J., Day, J. M. D., Nesvorny, D. & Elkins-Tanton, L. Stochastic late accretion to Earth, the Moon, and Mars. *Science* **330,** 1527–1530 (2010).
11. Raymond, S. N., Schlichting, H. E., Hersant, F. & Selsis, F. Dynamical and collisional constraints on a stochastic late veneer on the terrestrial planets. *Icarus* **226,** 671–681 (2013).
12. Touma, J. & Wisdom, J. Resonances in the early evolution of the Earth-Moon system. *Astron. J.* **115,** 1653–1663 (1998).
13. Ward, W. R. & Canup, R. M. Origin of the Moon's orbital inclination from resonant disk interactions. *Nature* **403,** 741–743 (2000).
14. Thompson, C. & Stevenson, D. J. Gravitational instability in two-phase disks and the origin of the Moon. *Astrophys. J.* **333,** 452–481 (1988).
15. Salmon, J. & Canup, R. M. Lunar accretion from a Roche-interior fluid disk. *Astrophys. J.* **760,** 83 (2012).
16. Wisdom, J. & Tian, Z. L. Early evolution of the Earth–Moon system with a fast-spinning Earth. *Icarus* **256,** 138–146 (2015).
17. Morbidelli, A., Marchi, S., Bottke, W. F. & Kring, D. A. A sawtooth-like timeline for the first billion years of lunar bombardment. *Earth Planet. Sci. Lett.* **355–356,** 144–151 (2012).
18. Canup, R. M. Dynamics of lunar formation. *Annu. Rev. Astron. Astrophys.* **42,** 441–475 (2004).
19. Nimmo, F. & Chen, E. M. A. Tidal dissipation in the early lunar magma ocean and its role in the evolution of the Earth-Moon system. *45th Lunar Planet. Sci. Conf.* abstr. 1459 (2014).
20. Alemi, A. & Stevenson, D. J. Why Venus has no moon. *Bull. Am. Astron. Soc.* **38,** 491 (2006).
21. Canup, R. M. Lunar-forming impacts: processes and alternatives. *Philos. Trans. R. Soc. London Ser. A* **372,** 20130175 (2014).
22. Spurzem, R., Giersz, M., Heggie, D. C. & Lin, D. N. C. Dynamics of planetary systems in star clusters. *Astrophys. J.* **697,** 458–482 (2009).

Acknowledgements This research was carried out as part of a Henri Poincaré Fellowship at the Observatoire de la Côte d'Azur (OCA) to K.P. The Henri Poincaré Fellowship is funded by the OCA and the City of Nice, France. A.M. thanks the European Research Council Advanced Grant ACCRETE (no. 290568).

Author Contributions K.P. and A.M. discussed every step of the project, designed the simulation set-up and co-wrote the numerical code. K.P. performed the simulations and the statistical analysis.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to K.P. (pahlevan@oca.eu).

METHODS

A large number of simulations (about 10^3) are required to characterize the distribution of outcomes for repeated encounters of a given planetesimal population with a given early Earth–Moon system. Accordingly, we design a numerical experiment that captures the physics of the problem statistically and that can be computed efficiently. Heliocentric orbits for late-accreting planetesimal populations were generated according to a Rayleigh distribution with a Rayleigh eccentricity $e_R = 0.3$ and inclination $i_R = e_R/2$; these values are consistent with simulations of terrestrial planet formation[23]. To test for sensitivity to population orbits, we varied e_R between 0.3 and 0.4; the resulting median inclination excitation changed by less than 10%. With given orbital distributions, the subset of the population that is Earth-crossing was selected and encounter probabilities with the Hill radius (R_H) of the Earth were calculated according to expressions given in ref. 24. The masses of the planetesimals are assumed to be in the range $0.15M_L$–$1.2M_L$ (M_L is the lunar mass), consistent with those expected for the projectiles carrying the Earth's late accretion[10]. At the beginning of the simulations, the Earth and the Moon were placed on circular uninclined orbits with radii of 1 AU and $5R_E$, respectively, near their orbits at the end of accretion and the beginning of tidal history. An encounter time and encounter orbit were chosen randomly according to the distribution of Earth-crossing planetesimal encounter probabilities. At the time of each encounter, phases for the lunar orbit, characterized by the argument of perigee (ω), the longitude of the ascending node (Ω) and the mean anomaly (M) were selected randomly, as was the orientation of the planetesimal orbit, within those orientations admitted by the selected orbital parameters. The impact parameter (b) was selected in the interval $[0, R_H]$ according to a uniform encounter probability per unit area ($dP \propto b\,db$). Gravitational three-body (Earth–Moon–planetesimal) encounters were integrated with a Bulirsch–Stoer integrator (included in the SWIFT package) in a geocentric reference frame, tracking changes to the lunar orbit. In between three-body encounters, the eccentricity (e) and semi-major axis (a) of the lunar orbit were evolved with a constant-Q tidal model[25] while the lunar inclination was evolved with a model[26] for planetary tides:

$$\frac{di}{i} = -\frac{1}{4}\frac{da}{a} \quad (1)$$

Impacts with the Earth were assumed to be inelastic merging events with the final body carrying the total mass and momentum. Impacts onto the Moon would have been in the erosive and/or catastrophic disruption regime and realizations with such events were removed from the subsequent analysis (discussed below).

Remnant planetesimals can, in general, be eliminated via collision with the terrestrial planets and the Sun or dynamical ejection from the inner Solar System[11]. We characterize such losses using the outcome of direct N-body simulations that trace the evolution of such early planetesimals, yielding a tenfold depletion of the Earth-crossing population in the first 100 Myr, which corresponds to a population decay law of approximately $\exp(-t/\tau_{SS})$, where t is time and τ_{SS} is the time constant for the decay of the population (approximately 44 Myr) (ref. 17). Although the modern near-Earth-object population is resupplied by the asteroid belt in a quasi-steady state fashion and effectively does not decay, the leftover planetesimal population is not resupplied by a larger reservoir and therefore does decay. Owing to partial resupply of Earth-crossing bodies, the timescale for the decay of the Earth-crossing population can nevertheless be different to the lifetime of individual particles. To integrate the decay rates of Earth-crossing planetesimals using our simulations, we use the following procedure. After generating orbital populations, but before running three-body integrations, we allow the Earth-crossing populations to encounter the Earth alone, which permits derivation of a time constant for decay of this population solely via collision with the Earth ($\tau_E = 79$ Myr). Next, we require that the Earth-crossing planetesimal population in our three-body simulations decay at the same average rate as that observed in the N-body heliocentric simulations. We therefore decompose average loss rates of the Earth-crossing population into terrestrial and non-terrestrial loss modes ($1/\tau_{SS} = 1/\tau_E + 1/\tau_{NE}$), and thereby derive a time constant ($\tau_{NE} = 99$ Myr) for removal via non-terrestrial loss channels. Accordingly, we stochastically remove bodies from the population in our three-body simulations such that the average loss rate of Earth-crossing planetesimals—through collision with Earth (explicit) as well as through other modes of loss (implicit)—is consistent with the average loss rates observed in N-body simulations of late accretion[17] (see Extended Data Fig. 1a).

Each data point in Fig. 2 is the result of 4,000 realizations. To analyse the results, certain realizations were eliminated. These realizations fall into three categories: those that result in (1) a collision of a planetesimal with the Moon; (2) dynamical loss of the Moon; and (3) too large or too small a mass accreted by the Earth.

(1) Occasionally, one of the planetesimals in our simulations collides with the Moon rather than with the Earth. For simulations that deliver the late-accreted mass to Earth via one, two and four planetesimals, the fraction of realizations in which such an event takes place is 9%, 15% and 25%, respectively. Given the masses that we assume for the planetesimals ($0.15M_L$–$1.2M_L$), it is doubtful that the Moon ever experienced such a massive collision. The largest lunar impact for which we have clear and unambiguous evidence is the South Pole Aitken Basin event (an approximately 10^{34} erg event[27]), which, at the encounter velocities considered here (see Extended Data Fig. 1b), corresponds to a lunar impactor that is 3–4 orders of magnitude less massive than the planetesimals in our populations. Although most basin-forming impacts are thought to have occurred during the late accretion era[17], the effect of these impacts on the lunar inclination was minor[18].

(2) Certain realizations, particularly those that correspond with the strongest tides and largest amount of interacting mass, generate excitations in eccentricity that are sufficient to destabilize the satellite orbit. Hence, collision with the host planet or (more rarely) liberation of the satellite into heliocentric orbit follows.

(3) The total amount of planetesimal mass at the start of simulations was chosen such that, on average, the mass accreted by the Earth would be $0.0075M_E$ or $0.015M_E$, hereafter denoted the 'target mass'. Given the stochasticity inherent to this problem, the accreted mass varies between realizations, resulting in a distribution of outcomes centred on the target mass. To facilitate the expression of the results in terms of late-accreted mass, we eliminate from the subsequent analysis those realizations whose accreted mass is greater than or less than the target mass.

Differential momentum transfer is the process underlying this excitation mechanism. For simplicity, we describe this process for the case of a planetesimal colliding with the Earth, but the general case of a collisionless encounter is similar. The orbit of the Earth and Moon around their common centre of mass is defined by their relative position and velocity. A third body encountering the Earth–Moon binary must have an orbit that crosses the system's heliocentric orbit, and approaches it with some finite velocity at large separation. Hence, the delivery of mass onto the Earth is accompanied by the delivery of external momentum that—in the impulse approximation—changes the relative velocity, but not the relative position, of the Earth with respect to the Moon, altering the mutual orbit, a hitherto overlooked effect that can excite the lunar inclination and eccentricity.

We ask whether the satellite excitation is dominated by the few strongest encounters or by the much more numerous distant ones. Theory suggests that for top-heavy perturber populations, the few strongest perturbations dominate over the more numerous weak ones[28]. We test this theory for our simulations by generating realizations where the impact parameter is chosen in the interval $[0, R_H]$, $[0, R_H/2]$ and $[0, R_H/4]$, progressively eliminating a large number of distant encounters. The results of the three simulations are statistically indistinguishable (see Extended Data Fig. 2a), confirming the theoretical expectation. To facilitate the reproducibility of our results, we plot a measure of the strength of a perturbation against the impact parameter of the encounters for two individual simulations (Extended Data Fig. 2b, c).

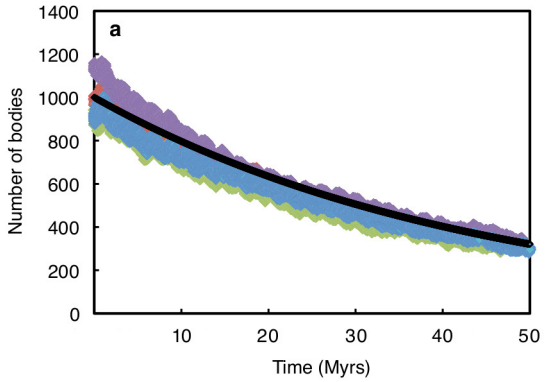
The simulations are permitted to proceed until the population of Earth-crossing bodies are exhausted either through collision with the Earth or through loss via another channel in accordance with an average rate (see above). At the end of the simulations, which characteristically last about 10^8 years, the lunar orbit is typically at about $40R_E$. To compare the simulation outcomes to the modern system, we propagate the lunar orbit forward to its current separation at $60R_E$ and permit the inclination to decay in accordance with the action of planetary tides (equation (1)). The number of simulations was chosen such that the median inclination values vary by only several per cent.

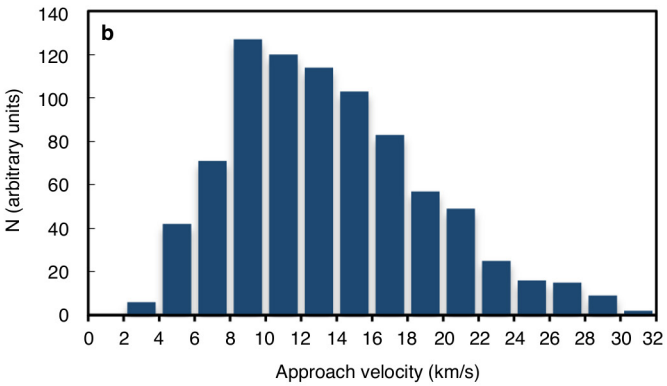
Recently, it has been suggested that the lunar inclination could have been damped via obliquity tides in the lunar magma ocean (LMO) as the Moon's obliquity increased during its approach to the Cassini state transition between $20R_E$ and $30R_E$ (ref. 19). The authors of ref. 19 put forward one interpretation of the current excited state of the lunar inclination: that the inclination was excited early[12,13], but that the rate of tidal dissipation in the post-giant-impact Earth was sufficiently low to delay passage of the lunar orbit through the Cassini state transition until after the crystallization of the LMO, at which point the effect of obliquity tides on the lunar inclination becomes much less. With the 'late' mechanism of lunar orbital excitation described here, we identify a different solution: that the rate of tidal dissipation on the post-impact Earth is sufficiently rapid in the first tens of millions of years to carry the Moon through the Cassini state transition and to damp any early acquired lunar inclination. Following such a resetting episode, the LMO crystallizes, and inclination excitation due to encounters is subsequently preserved. To explore this solution, we ran a suite of simulations in which the inclination is reset to zero until a certain time, and permitted to accumulate excitation subsequently (see Extended Data Fig. 3). Such a transition marks the time of crystallization of the LMO. It can be seen

that as long the duration of the LMO crystallization is sufficiently short, such a solution is viable and, indeed, necessary in a tidal evolution scenario recently described[29].

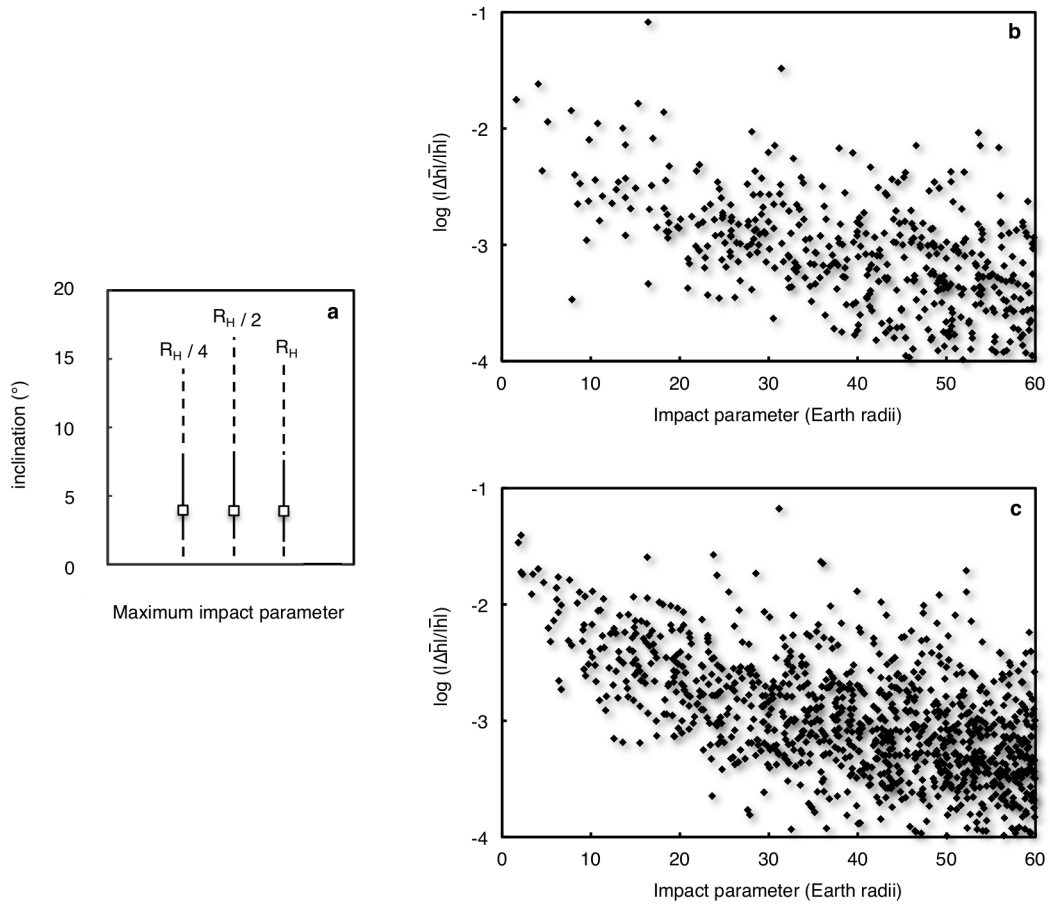
Code availability. The code used to conduct these simulations is available by request from the authors.

23. Walsh, K. J., Morbidelli, A., Raymond, S. N., O'Brien, D. P. & Mandell, A. M. A low mass for Mars from Jupiter's early gas-driven migration. *Nature* **475,** 206–209 (2011).
24. Wetherill, G. W. Collisions in the asteroid belt. *J. Geophys. Res.* **72,** 2429–2444 (1967).
25. Yoder, C. F. & Peale, S. J. The tides of Io. *Icarus* **47,** 1–35 (1981).
26. Kaula, W. M. Tidal dissipation by solid friction and the resulting orbital evolution. *Rev. Geophys.* **2,** 661–685 (1964).
27. Wieczorek, M. A., Weiss, B. P. & Stewart, S. T. An impactor origin for lunar magnetic anomalies. *Science* **335,** 1212–1215 (2012).
28. Collins, B. F. & Sari, R. Levy flights of binary orbits due to impulsive encounters. *Astron. J.* **136,** 2552–2562 (2008).
29. Zahnle, K. J., Lupu, R., Dobrovolskis, A. & Sleep, N. H. The tethered Moon. *Earth Planet. Sci. Lett.* **427,** 74–82 (2015).
30. Marchi, S., Bottke, W. F., Kring, D. A. & Morbidelli, A. The onset of the lunar cataclysm as recorded in its ancient crater populations. *Earth Planet. Sci. Lett.* **325–326,** 27–38 (2012).

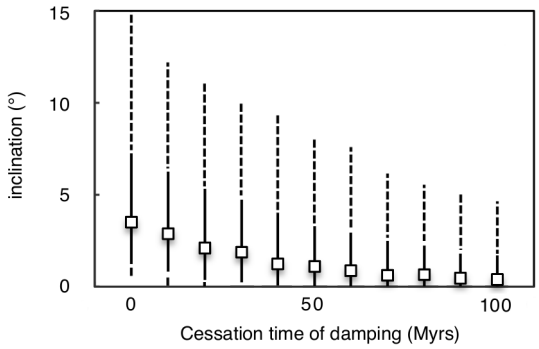




Extended Data Figure 1 | Properties of planetesimal populations. **a**, Decay rates of Earth-crossing planetesimal populations according to *N*-body simulations of the inner Solar System with a resonant (3:2) Jupiter and Saturn at 5.4 AU and 7.2 AU, respectively. Different colours represent the number of Earth-crossing bodies (hence the evolution is not monotonic) in different simulations, from recent integrations[17]. The black line is the prescribed decay rate used in the three-body simulations ($\tau_{SS} = 44$ Myr) to match the decay rate in heliocentric simulations. **b**, Histogram of implemented approach velocities (before acceleration due to Earth gravity) for late-accreting populations in three-body simulations. The population is generated using a Rayleigh distribution of eccentricities and inclinations ($e_R = 0.3$, $i_R = e_R/2$) and a semi-major axis range ($a = 0.8$–1.4 AU) that produces Earth-crossing orbits. These parameters are motivated by simulations of terrestrial planet formation, but the peak of the distribution (9 km s^{-1}) corresponds to the typical encounter velocity inferred for lunar basin-forming impactors[30].



Extended Data Figure 2 | Tests and outcomes for reproducibility. a, Test of the cumulative effect of repeated encounters: median values (squares), and 1σ (solid lines) and 2σ (dashed lines) intervals for three suites of simulations with $0.0075M_E$ accreted via a single body onto an Earth with strong dissipation ($k_2/Q = 0.1$). Each suite of simulations consists of incoming planetesimals with impact parameters (b) ranging from 0 to R_H, $R_H/2$ and $R_H/4$ (as indicated by the values given above each simulation result). The statistical similarity of the resultant distributions shows that distant encounters have a far smaller effect on inclination excitation than do the rare and strong close encounters. **b**, **c**, Distribution of 'kicks' versus impact parameter (b) of encounters from two different realizations. The change in the angular momentum vector of the satellite ($\Delta\bar{h}$) owing to encounter torques is normalized to the magnitude of the orbital angular momentum before the encounter ($\bar{h}$). The planetesimals approaching the Earth–Moon system in this simulation have a mass of $0.0075M_E$. Approach velocities are selected from the distribution plotted in Extended Data Fig. 1b. The cumulative effects of encounters with $b > 60R_E$ are negligible and therefore neglected. Panel **b** (**c**) shows data from a realization that lasts 26.2 Myr (45.3 Myr) and results in a satellite with a final inclination of $i = 1.9°$ ($i = 8.8°$). The strength of tidal dissipation used here ($k_2/Q = 0.1$) quickly results in a satellite semi-major axis of $a = 30R_E$–$40R_E$.



Extended Data Figure 3 | Effect of partial damping due to LMO obliquity tides. Median values (squares), and 1σ (solid lines) and 2σ (dashed lines) intervals for several suites of partially damped simulations. These simulations consist of an accreted mass of $0.0075M_E$ delivered via a single body onto a strongly dissipative ($k_2/Q = 0.1$) Earth, with the orbital excitation continuously reset ($e = 0$, $i = 0$) until a certain time and permitted to accumulate subsequently. Such a transition represents LMO crystallization and the cessation of inclination damping via obliquity tides[19]. These simulations are representative of excitation behaviour for partially damped cases. It can be seen that, so long as the crystallization of the LMO is sufficiently rapid (about 10^7 years), excitation via planetesimal encounters is largely unaffected.

LETTER

doi:10.1038/nature15768

Type-II Weyl semimetals

Alexey A. Soluyanov[1], Dominik Gresch[1], Zhijun Wang[2], QuanSheng Wu[1], Matthias Troyer[1], Xi Dai[3] & B. Andrei Bernevig[2]

Fermions—elementary particles such as electrons—are classified as Dirac, Majorana or Weyl. Majorana and Weyl fermions had not been observed experimentally until the recent discovery of condensed matter systems such as topological superconductors and semimetals, in which they arise as low-energy excitations[1–6]. Here we propose the existence of a previously overlooked type of Weyl fermion that emerges at the boundary between electron and hole pockets in a new phase of matter. This particle was missed by Weyl[7] because it breaks the stringent Lorentz symmetry in high-energy physics. Lorentz invariance, however, is not present in condensed matter physics, and by generalizing the Dirac equation, we find the new type of Weyl fermion. In particular, whereas Weyl semimetals—materials hosting Weyl fermions—were previously thought to have standard Weyl points with a point-like Fermi surface (which we refer to as type-I), we discover a type-II Weyl point, which is still a protected crossing, but appears at the contact of electron and hole pockets in type-II Weyl semimetals. We predict that WTe_2 is an example of a topological semimetal hosting the new particle as a low-energy excitation around such a type-II Weyl point. The existence of type-II Weyl points in WTe_2 means that many of its physical properties are very different to those of standard Weyl semimetals with point-like Fermi surfaces.

The band structure of some metals has non-trivial topological features[2]. Of such metals, the ones with vanishingly small density of states at the Fermi level—semimetals—stand out. For these materials, a distinction between topologically protected surface states and bulk metallic states can be made and their Fermi surfaces can be topologically characterized, unlike the case for metals, which have many states at the Fermi level. Two kinds of topological semimetals have attracted special attention: Dirac and Weyl semimetals. In these materials, a linear crossing of two (Weyl) or four (Dirac) bands occurs at the Fermi level (see Fig. 1a). The effective Hamiltonian for these crossings is given by the Weyl or gapless-Dirac equation, respectively. The Weyl crossings are protected from gapping, owing to the massless nature of the Weyl fermion. In the following, we limit the discussion to Weyl crossings only, although our results also hold for Dirac crossings.

The appearance of Weyl points (WPs) is possible only if the product of parity and time reversal is not a symmetry of the structure. When present, a WP acts as a topological charge—either a source or a sink of Berry curvature. A Fermi surface enclosing a WP has a well-defined Chern number, corresponding to the topological charge of this WP. Because the net charge must vanish in the entire Brillouin zone, WPs always come in pairs; they are stable to weak perturbations and are annihilated only in pairs of opposite charge. A large number of unusual physical phenomena are associated with Weyl topological semimetals, including the existence of open Fermi arcs in the surface Fermi surface[1,8] and various magnetotransport anomalies[9–15].

Weyl semimetals with broken time-reversal symmetry have been predicted to exist in several materials[1,16,17], but these predictions have yet to be experimentally verified. More recently, the Weyl semimetal was predicted to exist in inversion-breaking single-crystal nonmagnetic materials of the TaAs class[3,4]; this prediction has since been verified experimentally[5,6].

Weyl semimetals were previously thought to have a point-like Fermi surface at the WP. We refer to these as type-I WPs (WP1s), to distinguish them from the new type-II WPs (WP2s) that exist at the boundaries between electron and hole pockets, as illustrated in Fig. 1b. We discuss general conditions for WP2s to appear, and present evidence that WTe_2—the material with the largest never-saturating magnetoresistance reported[18] so far—is an example of the new type of topological semimetal hosting eight WP2s. These WP2s come in two quartets located 0.052 eV and 0.058 eV above the Fermi level. We present topological arguments that prove the existence of the new topological semimetal phase in WTe_2. We provide evidence of doping-driven topological Lifshitz transitions, which are characteristic of WP2s, as well as emerging Fermi arcs in the surface Fermi surface.

We start by considering the most general Hamiltonian describing a WP

$$H(\boldsymbol{k}) = \sum_{\substack{i=x,y,z \\ j=0,x,y,z}} k_i A_{ij} \sigma_j$$

where $\boldsymbol{k}$ is the wave vector in reciprocal space (crystal momentum vector), A is a 3×4 matrix of coefficients, σ_0 is the 2×2 unit matrix and and $\sigma_j, j = x, y, z$ are the three Pauli matrices. The energy spectrum is

$$\varepsilon_{\pm}(\boldsymbol{k}) = \sum_{i=x,y,z} k_i A_{i0} \pm \sqrt{\sum_{j=x,y,z} \left(\sum_{i=x,y,z} k_i A_{ij} \right)^2} = T(\boldsymbol{k}) \pm U(\boldsymbol{k}) \quad (1)$$

where $T(\boldsymbol{k})$ and $U(\boldsymbol{k})$ can be considered as the kinetic and potential components of the energy spectrum. $T(\boldsymbol{k})$, which is linear in momentum, tilts the cone-like spectrum $\varepsilon_{\pm}(\boldsymbol{k})$. This tilt breaks the Lorentz invariance of Weyl fermions in quantum field theory, but was previously considered unimportant. However, because Lorentz invariance does not need to be respected in condensed matter, its inclusion is important and leads to a finer classification of distinct Fermi surfaces, in correspondence with the theory of quadric surfaces, which suggests that there are exactly two distinct types of WPs (see Supplementary Information).

If, for a particular direction in reciprocal space, T is dominant over U, the tilt becomes large enough to cause a WP to appear at the point where the open electron and hole pockets touch, contrary to the standard case of a point-like Fermi surface. Thus, the condition for a WP to be of type II is that there exists a direction $\hat{\boldsymbol{k}}$ for which $T(\hat{\boldsymbol{k}}) > U(\hat{\boldsymbol{k}})$. If such a direction does not exist, then the WP is of type I. The clear qualitative distinction between the Fermi surfaces of the two types of WPs leads to marked differences in the thermodynamics of the hosting materials and their response to magnetic fields. In particular, in contrast to a WP1, which exhibits a chiral anomaly[9] for any direction of the magnetic field, the chiral anomaly appears in a WP2 only when the direction of the magnetic field is within a cone where $|T(\boldsymbol{k})| > |U(\boldsymbol{k})|$. If the field direction is outside this cone, then the Landau-level spectrum is gapped and has no chiral zero mode (see Supplementary Information).

On the lattice, the 'no-go' theorem[19] guarantees that Weyl fermions appear in pairs with Chern numbers of opposite sign. Because the

[1]Theoretische Physik and Station Q Zurich, ETH Zurich, 8093 Zurich, Switzerland. [2]Department of Physics, Princeton University, Princeton, New Jersey 08544, USA. [3]Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China.

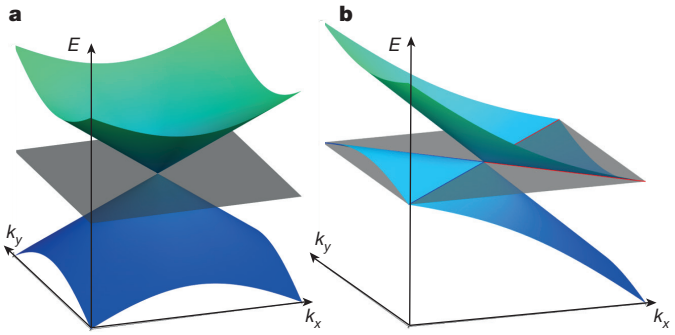


Figure 1 | Possible types of Weyl semimetals. a, Type-I WP with a point-like Fermi surface. **b**, A type-II WP appears as the contact point between electron and hole pockets. The grey plane corresponds to the position of the Fermi level, and the blue (red) lines mark the boundaries of the hole (electron) pockets.

Chern number of a WP is not changed by $T(\boldsymbol{k})$, WPs of different type can be chiral/anti-chiral partners of each other. The number of WPs of a certain type can be odd, but the total number of WPs must be even (for example, there can be one WP1 and one WP2).

We now describe WTe_2, a material we identified to host the new WPs. The crystal structure of WTe_2 is orthorhombic with space group $Pmn2_1$ (C_{2v}^7). Its primitive unit cell contains four formula units. The atomic structure is layered, with single layers of W separated from each other by Te bilayers and stacked along the z axis (see Supplementary Information). The distance between adjacent W atoms is considerably smaller along the x axis than it is along the y or z axes, creating strong anisotropy. The unit cell has two reflection symmetries: a mirror in the y–z plane m_{yz} and a glide plane g_{xz} formed by a reflection in the x–z plane followed by a translation by (0.5, 0, 0.5). Combined, they form a non-symmorphic twofold rotation C_2 (that is, a twofold rotation that is combined with a translation by a fraction of a lattice constant), which is important in the following symmetry arguments.

The result of band-structure calculations (see Supplementary Information) without spin–orbit coupling (SOC) is shown in Fig. 2a along the Γ–X direction, where an intermediate point $\Sigma = (0.375, 0, 0)$ is introduced. In addition to electron and hole pockets, 16 WPs per spin are found in WTe_2 in the absence of SOC (not shown in Fig. 2a). Half of these points occur at points of low symmetry with $k_z \neq 0$; the other half appear in the $k_z = 0$ plane, where the product of time reversal and C_2 ($C_{2T} = C_2T$) forms a little group. Generically, degeneracies on high-symmetry planes are forbidden; however, owing to the C_{2T} symmetry, twofold degeneracies are locally stable at points in the $k_z = 0$ plane. On the Γ–X line, the spectrum is generally gapped with a band-gap of approximately 1 meV, separating valence and conduction bands; see Fig. 2a.

Accounting for spin, but without SOC, bands become doubly degenerate, owing to opposite spin projections. This degeneracy doubles the topological charge of each WP because, by SU(2) symmetry, WPs corresponding to opposite spins have identical topological charge. Infinitesimal SOC cannot gap these WPs, giving a general criterion by which to search for Weyl semimetals: WPs are first found without SOC on the high-symmetry planes; the effects of SOC on these WPs are studied separately.

In WTe_2 SOC is not small. When turned on, it preserves electron and hole pockets, but substantially changes the structure of WPs. At intermediate SOC, WPs move, emerging or annihilating in pairs of opposite chirality. At full SOC, all WPs with $k_z \neq 0$ are annihilated. In the $k_z = 0$ plane, double degeneracies at isolated k points are still allowed by symmetry. Eight such gapless points are found, formed by the topmost valence and lowest conduction bands at full SOC. A pair of such points is shown in Fig. 2c. The other three pairs are related to this one by reflections. Energetically, both points are located only slightly (0.052 eV and 0.058 eV) above the Fermi energy E_F; see Supplementary Information for details.

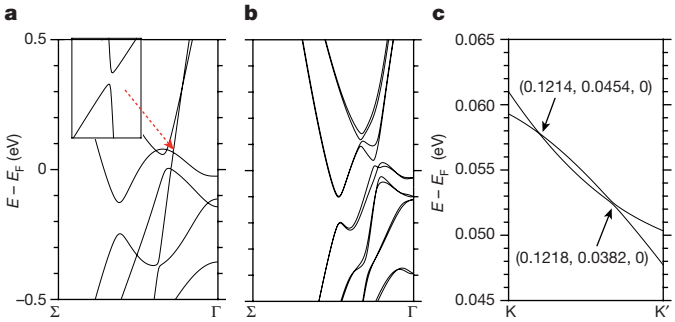


Figure 2 | Band structure of WTe_2. a, Band structure of WTe_2 without SOC. A fraction of the Γ–X segment is shown: the point Σ has coordinates (0.375, 0, 0). A bandgap of approximately 1 meV is shown in the inset, signalling a gapless point nearby. **b**, Band structure of WTe_2 with SOC. **c**, One of the four pairs of WPs is shown along the line K–K′, where K = (0.1208, 0.0562, 0) and K′ = (0.1226, 0.0238, 0). Their locations are designated in reduced coordinates (in units of reciprocal lattice constants).

Establishing degeneracies of bands (and the existence of WPs) computationally (or by inspection) is prone to finite-size effects: a point thought to be a degeneracy point might turn out to have a minuscule gap upon increasing computational precision. To rigorously establish the presence of WPs, we performed many tests that involve computing topological indices. The topological charge (±1) of each WP was found using an extension of the Wilson-loop and hybrid-Wannier-centres methods[20,21] to type-II Weyl semimetals. $\mathbb{Z}_2$ topological indices were also computed on several planes (including those in both standard and non-standard geometries) in the Brillouin zone. In total, these tests not only proved the existence of WPs, but also elucidated the structure of the Berry-flux connection between WPs and of the Fermi arcs on the surface of WTe_2. The resultant Fermi-arc structure is consistent with the calculations presented below. A detailed description of topological indices and ways to obtain them are found in Supplementary Information.

To check the nature of the WPs, we obtained the energy spectrum around them from first-principles calculations and fitted it to the theoretical model derived by symmetry analysis (Supplementary Information). Considering only linear terms in k_i—the momentum relative to the position of the WP—the spectrum in equation (1) becomes

$$\varepsilon_{\pm}(\boldsymbol{k}) = Ak_x + Bk_y \pm \sqrt{e^2k_z^2 + (ak_x + ck_y)^2 + (bk_x + dk_y)^2}$$

The values of the parameters A, B, a, b, c, d and e are given in the Supplementary Information. The kinetic component of the energy dominates along the line connecting this WP to its nearest neighbour (see Fig. 2c and Supplementary Information). We thus conclude that WTe_2 is a type-II Weyl semimetal.

We now discuss the Fermi surface topology and possible topological Lifshitz transitions in WTe_2. The evolution of the Fermi surface obtained from first-principles calculations is shown in Fig. 3 for different values of E_F. Owing to reflection symmetries, only part of the $k_z = 0$ plane of the Fermi surface is shown. For $E_F = 0$ eV, the Fermi surface is formed of two pairs of electron pockets and two pairs of hole pockets (eight pockets in total), which are separated in momentum space. For each pair, the larger pocket completely encloses the smaller one, in agreement with experiments[22]. This property is illustrated in Fig. 3a, where four halved pockets (two electron and two hole) are shown. The other halves are obtained by the glide reflection g_{xz}, and the remaining four pockets with $k_x > 0$ are obtained by the mirror reflection m_{yz}. All Fermi surfaces have zero Chern numbers when $E_F = 0$.

When E_F is raised, two additional electron pockets appear; the previously existing electron pockets persist. The hole pockets shrink quickly, two disappearing completely. Each of the remaining two split into two disconnected pockets. As a result, there are six electron pockets and four hole pockets in total (see Fig. 3b). When the Fermi level

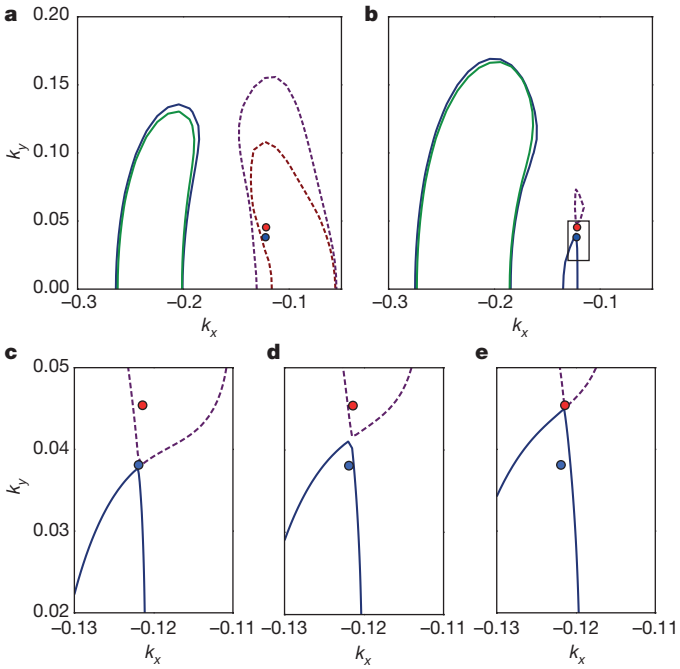


Figure 3 | Fermi surface at $k_z = 0$. A part of the Brillouin zone is shown. **a**, $E_F = 0$ eV; electron pockets (blue and green solid lines) and hole pockets (red and magenta dashed lines) come in pairs. WP with Chern number +1 (−1) is shown in red (blue). **b**, The representative structure of electron and hole pockets at higher energies ($E_F = 0.055$ eV shown). There are four hole pockets (one shown; dashed magenta line) and six electron pockets (halves of three of them shown; blue and green solid lines). The boxed region is the region shown in **c**–**e** for different values of E_F. **c**, $E_F = 0.052$ eV is set to the lower-energy WP. Contact between electron and hole pockets occurs at this WP. **d**, $E_F = 0.055$ eV is set to be between the two WPs. The electron and hole pockets are disconnected. The hole pocket encloses a WP with a Chern number $C = +1$. The electron pocket encloses the WP with $C = -1$ and its mirror image (not shown); the net Chern number of this pocket is zero. **e**, When $E_F = 0.058$ eV is set to the higher-energy WP, electron and hole pockets touch again (shown). They reopen at larger E_F with zero Chern numbers.

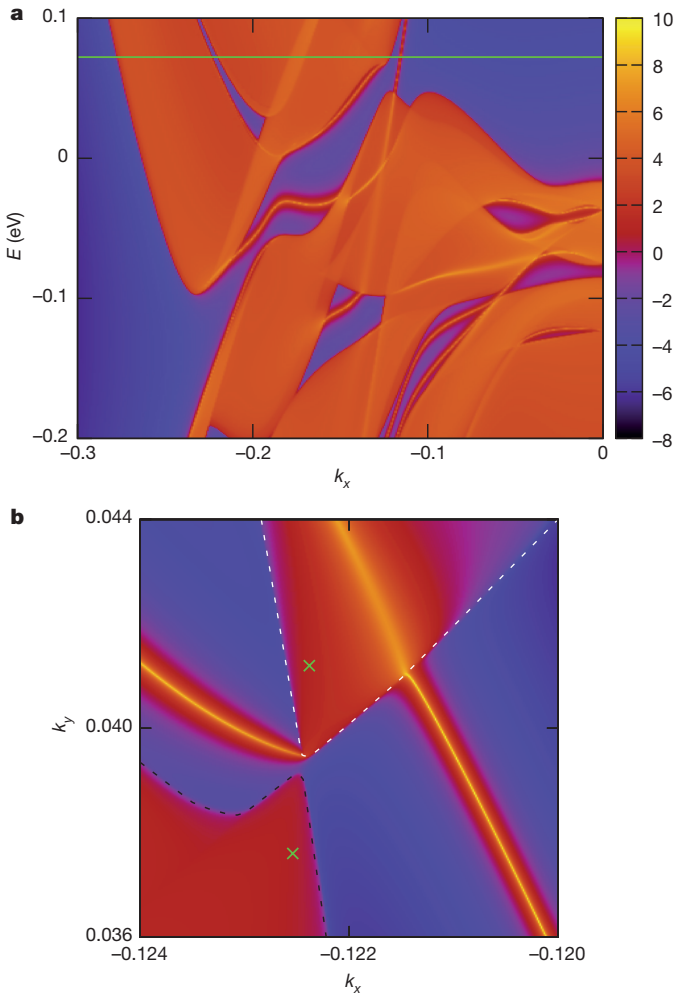


Figure 4 | Topological surface states. a, Spectral function of the (001) surface. The Fermi level (green line) is set to be between the WPs. **b**, Fermi surface of the (001) surface and a Fermi arc connecting hole and electron pockets. Green crosses mark the positions of WPs.

is tuned to the first WP, $E_F = 0.052$ eV (corresponding to the addition of approximately 0.064 electrons per unit cell), each of the two newly appeared electron pockets touches two hole pockets at the positions of the WPs, as illustrated in Fig. 3c for part of $k_z = 0$ plane. Further increase of E_F disconnects the electron and hole pockets again—see Fig. 3d for $E_F = 0.055$ eV—but with changed topology: electron pockets still have zero Chern numbers because they enclose two WPs of opposite charge, related by g_{xz}. The hole pockets have Chern numbers of ± 1. Topologies of the other hole pockets are obtained by changing the sign of the Chern number according to the appropriate mirror and glide symmetries. The pockets touch again (see Fig. 3e) when the Fermi level is tuned to the higher-energy WP, $E_F = 0.058$ eV (corresponding to approximately 0.079 additional electrons per unit cell). Upon raising E_F further, the pockets disconnect again, and all Fermi-surface Chern numbers become zero.

To facilitate the observation of topological Lifshitz transitions, hydrostatic pressure is applied. Neighbouring WPs are pushed away from each other in k space under compression. In particular, a 0.5% (2%) compression increases the distance between the WPs from 0.7% to 2% (4%) of the reciprocal vector $|\boldsymbol{G}_2|$ (see Supplementary Information for a discussion of strain effects, including how to obtain only four WPs).

Finally, we discuss the topological surface states of WTe_2. Owing to reflection symmetries, WPs of opposite chirality are projected on top of each other on the (100) and (010) surfaces, which hence do not exhibit topologically protected surface states.

For the (001) surface, all the WPs project onto distinct points; hence, topological surface states appear. When E_F is tuned to be between the WPs, the hole pocket has non-zero Chern number and a Fermi arc emerges from it, connecting it to the WP of opposite Chern number inside the electron pocket. Figure 4a illustrates the spectral function of the (001) surface, where surface states connecting electron and hole bands are clearly visible. The Fermi surface of this surface has a topological Fermi arc (Fig. 4b) connecting projections of the topological hole (Fig. 3c–e) and electron pockets. The other surface state crossing the hole pocket emerges from the electron pocket (not seen in Fig. 4) and goes back into it, and thus can be pushed into the continuum of bulk states (see Supplementary Information).

Of other transition metal dichalcogenides, another strong candidate material is $MoTe_2$ (ref. 23), which is reported to be a semimetal resembling pressurized WTe_2. This material can also be used to explore new physical phenomena arising in the new topological semimetal phase presented here.

Received 1 June; accepted 22 September 2015.

1. Wan, X., Turner, A. M., Vishwanath, A. & Savrasov, S. Y. Topological semimetal and Fermi-arc surface states in the electronic structure of pyrochlore iridates. *Phys. Rev. B* **83,** 205101 (2011).
2. Volovik, G. E. *The Universe in a Helium Droplet* (Oxford Univ. Press, 2009).
3. Weng, H., Fang, C., Fang, Z., Bernevig, B. A. & Dai, X. Weyl semimetal phase in noncentrosymmetric transition-metal monophosphides. *Phys. Rev. X* **5,** 011029 (2015).
4. Huang, S.-M. *et al.* An inversion breaking Weyl semimetal state in the TaAs material class. *Nature Commun.* **6,** 7373 (2015).
5. Xu, S.-Y. *et al.* Discovery of a Weyl fermion semimetal and topological Fermi arcs. *Science* **349,** 613–617 (2015).

6. Lv, B. Q. *et al.* Experimental discovery of Weyl semimetal TaAs. *Phys. Rev. X* **5,** 031013 (2015).
7. Weyl, H. Elektron und Gravitation. *I. Z. Phys.* **56,** 330–352 (1929).
8. Silaev, M. A. & Volovik, G. E. Topological Fermi arcs in superfluid ^{3}He. *Phys. Rev. B* **86,** 214511 (2012).
9. Nielsen, H. B. & Ninomiya, M. The Adler–Bell–Jackiw anomaly and Weyl fermions in a crystal. *Phys. Lett. B* **130,** 389–396 (1983).
10. Zyuzin, A. A. & Burkov, A. A. Topological response in Weyl semimetals and the chiral anomaly. *Phys. Rev. B* **86,** 115133 (2012).
11. Hosur, P. & Qi, X. Recent developments in transport phenomena in Weyl semimetals. *C. R. Phys.* **14,** 857–870 (2013).
12. Volovik, G. E. Kopnin force and chiral anomaly. *JETP Lett.* **98,** 753–757 (2014).
13. Zhang, C. *et al.* Observation of the Adler–Bell–Jackiw chiral anomaly in a Weyl semimetal. Preprint at http://arXiv.org/abs/1503.02630 (2015).
14. Xiong, J. *et al.* Signature of the chiral anomaly in a Dirac semimetal: a current plume steered by a magnetic field. Preprint at http://arXiv.org/abs/1503.08179 (2015).
15. Huang, X. *et al.* Observation of the chiral-anomaly-induced negative magnetoresistance in 3D Weyl semimetal TaAs. *Phys. Rev. X* **5,** 031023 (2015).
16. Xu, G., Weng, H., Wang, Z., Dai, X. & Fang, Z. Chern semimetal and the quantized anomalous Hall effect in $HgCr_2Se_4$. *Phys. Rev. Lett.* **107,** 186806 (2011).
17. Burkov, A. A. & Balents, L. Weyl semimetal in a topological insulator multilayer. *Phys. Rev. Lett.* **107,** 127205 (2011).
18. Ali, M. N. *et al.* Large, non-saturating magnetoresistance in WTe_2. *Nature* **514,** 205–208 (2014).
19. Nielsen, H. B. & Ninomiya, M. Absence of neutrinos on a lattice: (I). Proof by homotopy theory. *Nucl. Phys. B* **185,** 20–40 (1981).
20. Soluyanov, A. A. & Vanderbilt, D. Computing topological invariants without inversion symmetry. *Phys. Rev. B* **83,** 235401 (2011).
21. Yu, R., Qi, X. L., Bernevig, A., Fang, Z. & Dai, X. Equivalent expression of $\mathbb{Z}_2$ topological invariant for band insulators using the non-Abelian Berry connection. *Phys. Rev. B* **84,** 075119 (2011).
22. Pletikosić, I., Ali, M. N., Fedorov, A. V., Cava, R. J. & Valla, T. Electronic structure basis for the extraordinary magnetoresistance in WTe_2. *Phys. Rev. Lett.* **113,** 216601 (2014).
23. Brown, B. E. The crystal structures of WTe_2 and high-temperature $MoTe_2$. *Acta Crystallogr.* **20,** 268–274 (1966).

Supplementary Information is available in the online version of the paper.

Acknowledgements A.A.S., D.G., QS.W. and M.T. acknowledge the support of Microsoft Research, the Swiss National Science Foundation through the National Competence Center in Research MARVEL and the European Research Council through ERC Advanced Grant SIMCOFE. Z.W. and B.A.B. acknowledge the support of ARO MURI W911NF-12-1-0461, ONR-N00014-11-1-0635, NSF CAREER DMR-0952428, NSF MRSEC DMR-0819860, the Packard Foundation and a Keck grant. X.D. is supported by the National Natural Science Foundation of China, the 973 program of China (no. 2011CBA00108 and no. 2013CB921700) and the "Strategic Priority Research Program (B)" of the Chinese Academy of Sciences (no. XDB07020100).

Author Contributions All authors contributed to performing the calculations and the analysis of the results.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.A.S. (soluyanov@itp.phys.ethz.ch).

LETTER

doi:10.1038/nature16066

Ultrafast ultrasound localization microscopy for deep super-resolution vascular imaging

Claudia Errico[1,2,3], Juliette Pierre[1,2,3], Sophie Pezet[4,5], Yann Desailly[1,2,3], Zsolt Lenkei[4,5], Olivier Couture[1,2,3]* & Mickael Tanter[1,2,3]*

Non-invasive imaging deep into organs at microscopic scales remains an open quest in biomedical imaging. Although optical microscopy is still limited to surface imaging owing to optical wave diffusion and fast decorrelation in tissue, revolutionary approaches such as fluorescence photo-activated localization microscopy led to a striking increase in resolution by more than an order of magnitude in the last decade[1]. In contrast with optics, ultrasonic waves propagate deep into organs without losing their coherence and are much less affected by *in vivo* decorrelation processes. However, their resolution is impeded by the fundamental limits of diffraction, which impose a long-standing trade-off between resolution and penetration. This limits clinical and preclinical ultrasound imaging to a sub-millimetre scale. Here we demonstrate *in vivo* that ultrasound imaging at ultrafast frame rates (more than 500 frames per second) provides an analogue to optical localization microscopy by capturing the transient signal decorrelation of contrast agents—inert gas microbubbles. Ultrafast ultrasound localization microscopy allowed both non-invasive sub-wavelength structural imaging and haemodynamic quantification of rodent cerebral microvessels (less than ten micrometres in diameter) more than ten millimetres below the tissue surface, leading to transcranial whole-brain imaging within short acquisition times (tens of seconds). After intravenous injection, single echoes from individual microbubbles were detected through ultrafast imaging. Their localization, not limited by diffraction, was accumulated over 75,000 images, yielding 1,000,000 events per coronal plane and statistically independent pixels of ten micrometres in size. Precise temporal tracking of microbubble positions allowed us to extract accurately in-plane velocities of the blood flow with a large dynamic range (from one millimetre per second to several centimetres per second). These results pave the way for deep non-invasive microscopy in animals and humans using ultrasound. We anticipate that ultrafast ultrasound localization microscopy may become an invaluable tool for the fundamental understanding and diagnostics of various disease processes that modify the microvascular blood flow, such as cancer, stroke and arteriosclerosis.

The recent discovery of super-resolution optical microscopy led to a revolutionary improvement of resolution through the use of different technical approaches[1,2]. One major implementation, fluorescence photo-activated localization microscopy (FPALM), exploits the stochastic blinking of specific fluorescent sources to separate them into individual events in independent frames. A super-resolved image is obtained by localizing the centre of each separable source and accumulating these positions over thousands of acquisitions. The resulting image highlights structures that are hundreds of times smaller than the wavelength, such as the cell membrane and small organelles[3].

In clinical ultrasound imaging, intravenously injected contrast agents (1–3-μm-diameter microbubbles) act as intravascular acoustic sources to reveal the vascular bed. At typical concentrations, a cloud of microbubbles can be considered as a sub-wavelength random distribution of Rayleigh scatters. The resolution of ultrasound contrast imaging is limited by the classical wave diffraction theory and corresponds roughly to the ultrasonic wavelength (typically between 200 μm and 1 mm in clinical applications). Nevertheless, thanks to the advent of ultrafast ultrasound imaging[4], we recently proposed an ultrasound equivalent of FPALM[5,6] that surpassed the conventional diffraction limit of echography by more than tenfold. The use of ultrafast acquisitions based on plane wave transmissions at the rate of a thousand frames per second may lead to several key advantages when imaging contrast agents. First, the decorrelation of the microbubble signal from frame to frame is typically in the millisecond range[7]. As the tissue signature decorrelates more slowly than the microbubble signal, it is thus removed by simply applying a differential subtraction filter of consecutive frames. Second, since they respond to ultrasound differently over several frames, microbubbles blink separately through the spatiotemporal differentiation process and become temporally separable sources. Last, since the ultrasonic sequence provides simultaneously very high temporal resolution in all pixels of the image, it becomes possible to track the signature of many individual microbubbles both in space and time and thus to quantify the local blood flow speed over a very large dynamical range. As ultrasonic waves can penetrate several centimetres of tissue, extracting the positions of each of these bubbles could lead to the full reconstruction of the deep vascular system down to the level of capillaries. However, the usefulness of these theoretical benefits remains to be demonstrated *in vivo*.

Current methods for *in vivo* microvascular imaging are limited by trade-offs between the depth of penetration, resolution and acquisition time. For instance, microcomputed tomography[8] and magnetic resonance imaging[9] are able to resolve vessels down to a few tens of micrometres with deep tissue penetration, but they remain limited by long scanning times. Near-infrared II fluorescence imaging[10] has high spatial resolution (~50 μm) and fast acquisition times (<200 ms). Nevertheless, it lacks sufficient tissue penetration (<1–3 mm) for whole-brain imaging. High-resolution photoacoustic imaging[11] does not require contrast agents and can attain resolutions of a few micrometres, but also lacks penetration (0.75 mm). Finally, acoustic angiography resolves tumour vessels around 150 μm in diameter, but is still hampered by the trade-off between penetration and resolution[12].

Here, we demonstrate ultrafast ultrasound localization microscopy (uULM), which combines deep penetration and super-resolution imaging at unprecedented spatiotemporal resolution, by using clinically approved contrast agents: inert gas microbubbles. uULM is implemented *in vivo* on anaesthetized male Sprague–Dawley rats fixed within a stereotactic frame. Their skull was either left intact or thinned to reduce the acoustic attenuation caused by the bone. We used a small

[1]INSERM, Institut Langevin, 1 rue Jussieu, 75005 Paris, France. [2]Institut Langevin, ESPCI-ParisTech, PSL Research University, 1 rue Jussieu, 75005 Paris, France. [3]CNRS UMR 7587, 1 rue Jussieu, 75005 Paris, France. [4]CNRS, UMR 8249, 10 rue Vauquelin, 75005 Paris, France. [5]Brain Plasticity Unit, ESPCI-ParisTech, PSL Research University, 10 rue Vauquelin, 75005 Paris, France.
*These authors contributed equally to this work.

ultrasonic probe, connected to a fully programmable ultrafast ultrasound scanner to image a coronal slice of the brain.

The major challenge of ULM is to intercept a sufficient number of separable sources (microbubbles) in the blood stream to obtain super-resolved vasculature maps over a large region within a reasonable acquisition time. Therefore, we detected microbubbles in the rat brain cortex by looking at their fast decorrelation within a stack of 75,000 images acquired continuously for 150 s. The millisecond-timescale decorrelation of the microbubble signal can be generated by several processes, including disruption, dissolution and motion. In the current implementation, pulse sequences were chosen to reduce ultrasound-induced disruption or dissolution of microbubbles. As microbubbles are point-scatters and since small variations of phase can be detected in the radio-frequency data, microbubble displacement much smaller than the wavelength appears as a strong decorrelation signal on differential filtered images. Moreover, by exploiting the coherence of backscattered signals, the spatiotemporal filtering approach discriminates slowly moving objects of sub-wavelength size (low spatial coherence), that is, bubbles, from slow motion tissue signals whose temporal variations affect many neighbouring pixels the same way (high spatial coherence). The ultrafast frame rate was achieved by emitting plane waves and collecting the backscattered echoes with all the array elements. For each transmission, the resulting echoes were exploited to reconstruct *in silico* an entire ultrasonic frame by using parallel beamforming. In the averaged stack of ultrasound images only the thinned skull was observable (Fig. 1a). The decorrelation of bubbles was detected using frame-to-frame differential processing, which yields individual and fast-changing sources within the ultrafast ultrasound images (Fig. 1b). This high-pass filter uses the very high spatiotemporal sampling to eliminate tissue and skull signals. Since microbubbles are much smaller than the wavelength (1–3 μm versus 100 μm) and can be individually separated in space and time, they appeared as the point-spread function (PSF) of the ultrasound system. The spatial coordinates of the bubble centroids were extracted one by one by deconvolving the individual sources from the predicted Gaussian PSF. As these sources are locally unique, each of these positions can be estimated with a 2.5 μm maximum theoretical resolution in the axial direction. For example, a blinking microbubble flowing in vessels at the level of the primary somatosensory forelimb or hindlimb cortex (S1HL/FL), appeared as a spot representing the centre of the interpolated PSF (Fig. 1c).

Typically, we localized in 150 s about 1,000,000 events within one hemisphere of the brain cortex. Furthermore, we were able to track each moving bubble according to its instantaneous position and in-plane velocity vector, leading to quantitative and localized maps of cerebral blood flow velocity. Hence, ultrafast imaging allows the reconstruction of entire organs within tens of seconds, a prerequisite for a preclinical and clinical modality. Far beyond a technological leap, ultrafast imaging ensures the necessary discrimination between single bubble signatures and tissue at high bubble concentrations using optimal spatiotemporal clutter filters[13]. By tracking the local motion of bubbles at a kHz rate, it estimates their motion over a very large dynamic range of velocities and consequently vessel diameters (1 mm s^{-1} to several cm s^{-1} and 15 μm to 1–5 mm, respectively) during a sufficiently long acquisition time simultaneously in all voxels of the image. Finally, in fast-moving or pulsatile organs, tissue motion correction could be assessed through speckle tracking with micrometric sensitivity to co-register bubble positions in real time or post-processing[5,14,15]. This remains a fundamental asset with respect to individual bubble localization techniques based on conventional ultrasound sequences, recently[16] discussed, which need to separate echoes through high dilution of contrast agents and image clamped tissue for extended durations (1 h) because of limited frame rates[17].

We obtained extremely detailed structural reconstructions of the microvasculature in the rat brain cortex (5 mm width and 3 mm

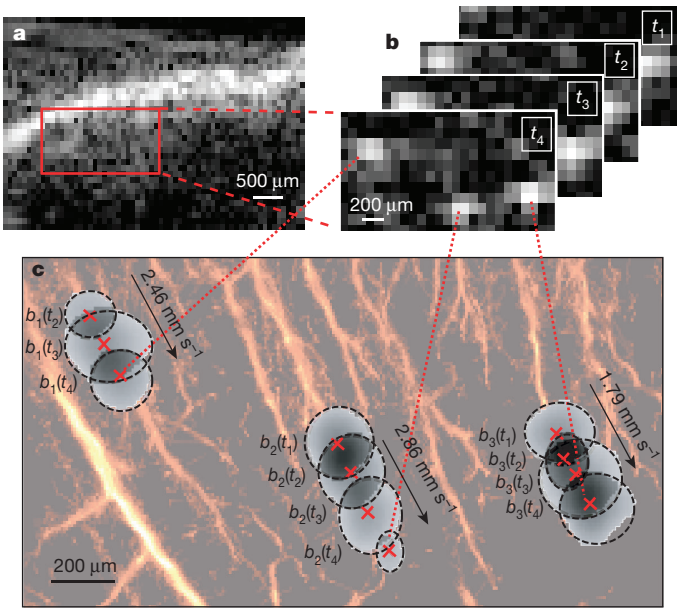


Figure 1 | Principle of uULM. a, Ultrafast detection of individual sources from a low-quality B-mode image (averaged stack of 250 beamformed images), through a thinned skull. **b**, Four representative frames were separated by 44 ms (t_1–t_4) and filtered to remove the slow-moving tissue signal. **c**, Three independent microbubbles blinking over several milliseconds from **b** were followed in the region of interest within the cortex. The echo of each bubble event (high-contrast pixels) was deconvolved with the PSF to obtain the exact position of the centroid (red crosses). Superposition of thousands of occurrences yields a highly resolved localization map for this region.

depth) under the thinned skull window (Fig. 2a), displaying vessels with diameters between 15 μm and 65 μm. The images were reconstructed with a pixel size of 10 μm × 8 μm, corresponding to a tenfold increase in resolution as compared to conventional ultrasound imaging. Furthermore, bifurcations of the penetrating arterioles within the S1HL/FL were easily observable down to the terminal branching points (Fig. 2a), where vessels attain the hypovascular white matter[18]. In comparison, the contrast-enhanced image created using conventional power Doppler is limited by diffraction (Fig. 2b)[19], highlighting only the large vessels of the rat brain cortex without distinguishing details below the wavelength scale. Moreover, Doppler detection is strongly biased towards flows that are perpendicular to the array.

More detailed analysis of the cross-section of individual vessel profiles, indicated by lines 1 and 2 in Fig. 2a, yielded diameter sizes of 17 μm and 9 μm full-width at half-maximum, respectively, corresponding to capillaries[20] (Fig. 2c). These values represent a convolution between the actual size of the vessel and the response of the localization microscopy method, giving an upper limit to its resolution (wavelength $\lambda/10$). Investigation of a branching vessel profile (profile 3 in Fig. 2a) showed that at a distance of 16 μm ($\lambda/6$), the two vessels are still clearly separated. Such high resolution depends on the number of bubbles present in the reconstructed pixel (10 μm × 8 μm) and could thus be further improved with longer integration times.

Next, we evaluated the ability of our method to measure blood flow dynamics in cortical microvessels. Measured blood flow in-plane velocities in the rat brain showed a large dynamic range up to several cm s^{-1} for large vessels and down to 2 mm s^{-1} in small vessels. Blood flow velocity inside of the relatively large penetrating artery was well resolved (profile 4 in Fig. 2d, e, profile 5 in Fig. 2d, f) and was inversely correlated with vessel diameter, showing 15 mm s^{-1} maximum velocity at 80 μm diameter and 2 mm s^{-1} maximum velocity at 15 μm diameter, consistent with the literature values[21,22]. Interestingly, it was clear that larger vessels support higher flow within their centre with respect

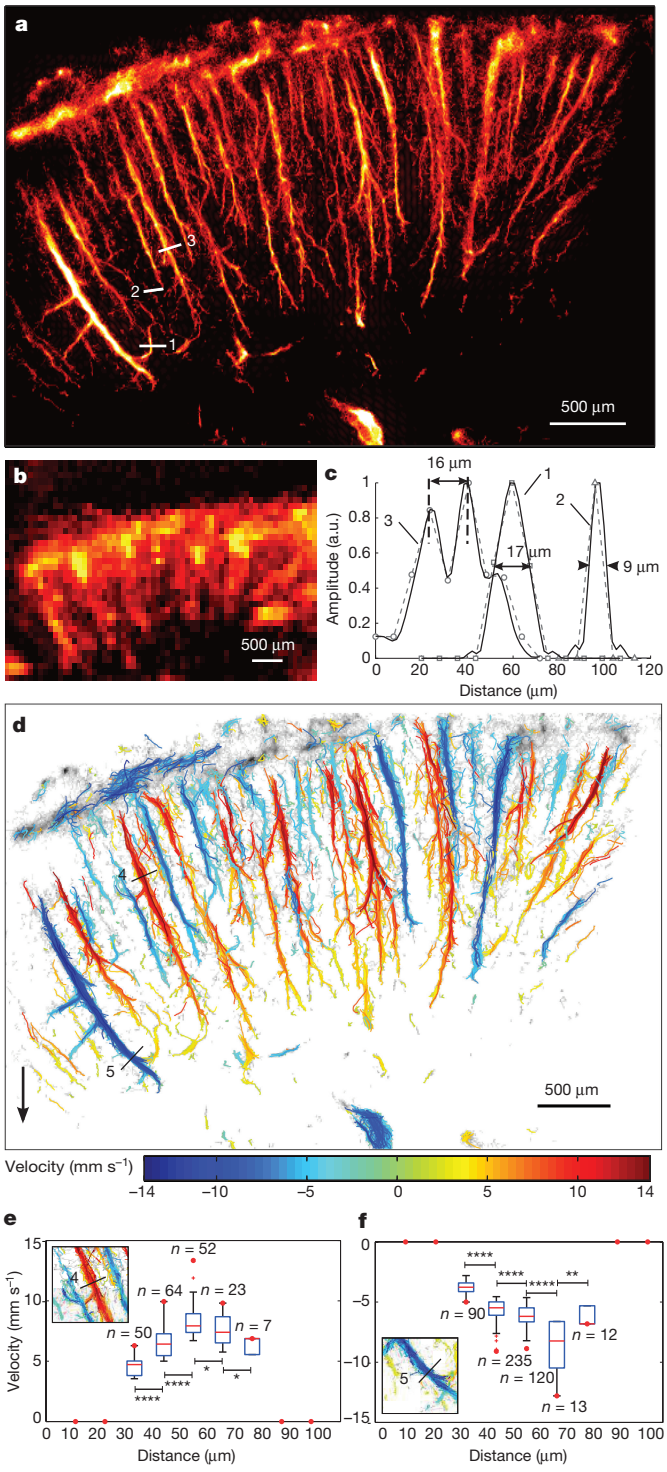


Figure 2 | Spatial resolution and quantification of uULM in the rat brain cortex through a thinned skull window. a, Microbubble density maps were reconstructed with a spatial resolution of $\lambda/10$ (pixel size = 8 µm × 10 µm). **b**, Same area in a conventional power Doppler image. **c**, Interpolated profiles along the lines marked in **a** display 9 µm vessels (2) and resolve two vessels closer than 16 µm (3). a.u., arbitrary units. **d**, Dynamic tracking of bubbles separates vessels in two populations with opposite blood flow direction. Positive values indicate blood flow distancing from the probe. Bubble velocities between 1 mm s^{-1} and 14 mm s^{-1} are detectable. **e**, **f**, Velocity profiles associated with lines 4 (**e**) and 5 (**f**) in **d**. Red line, median; blue box, 25th to 75th percentile; whiskers extend to the most extreme data points that are not considered outliers; other points, outliers. Unpaired Student's *t*-test. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$.

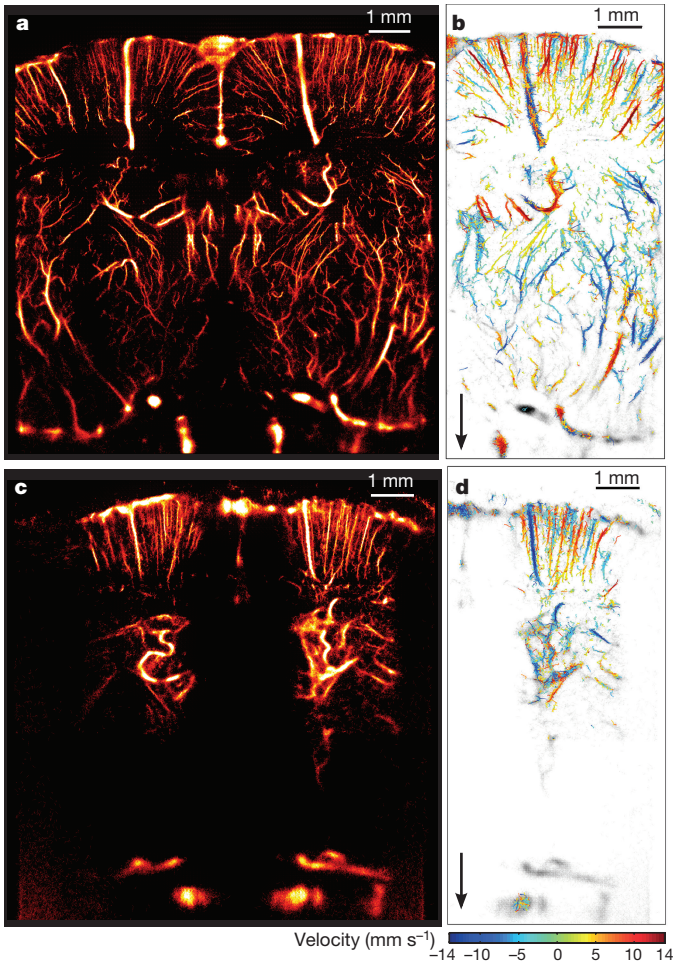


Figure 3 | uULM of the rat brain through a thinned skull window or through the intact skull. a, uULM performed through a thinned skull at a coronal section, Bregma −1.5 mm, providing a resolution of 10 µm × 8 µm in depth and lateral direction, respectively. **c**, uULM performed through the intact skull at Bregma −1 mm. Owing to the attenuation of the ultrasound waves in the presence of the bone, the achieved resolution was 12.5 µm × 1 µm in depth and lateral direction, respectively. Thus, the smallest vessel detectable was 20 µm wide. **b**, **d**, In-plane velocity maps from parts of the vessels in **a** and **c**, respectively.

to their periphery. Although the images are integrated over a slab of about 100 µm thick, we could separate two sets of vessels simply on the basis of their flow velocities. Some bubbles were travelling at a much slower speed in the opposite direction than the background venules. Moreover, in contrast with conventional ultrasound Doppler imaging, which is sensitive mostly to flow towards or away from the ultrasound probe, here we also observed and measured microbubbles that were moving sideways. This is particularly useful to observe the tortuosity of the small vessels and detect abrupt branching in vessels within the cortex.

In-plane velocity measurements can define the resolution of uULM. We consider that two resolution cells are distinguishable if their velocity distributions are statistically different ($P < 0.05$). The median of the upper half of the velocity distribution for each resolution cell is displayed in Fig. 2e, f. When the resolution cells are 8–12 µm in size, adjacent pixels can be considered distinct. Interestingly, the maximum velocities follow a parabolic profile, as expected for vessels of this size.

Finally, we investigated the spatial coverage of our imaging method. At 15 MHz, the attenuation of an ultrasound wave within brain tissue is approximately 5 dB cm^{-1} (ref. 23), which allows imaging at several centimetres depth. Super-resolved images could be obtained

in vivo over the entire depth of the brain (12.5 mm at Bregma −1.5 mm; Fig. 3a), demonstrating that ULM can map vessels below the rat brain cortex over several coronal planes (Extended Data Fig. 2). Super-resolved imaging is also possible through the intact skull (Fig. 3c; Bregma −1.0 mm) but the lower signal-to-noise ratio, resulting from skull-induced signal attenuation, globally reduces the number of localized microbubbles, increasing the limit of the smallest detectable vessel. However, this non-invasive version of our imaging method can still detect vessels that are 20 μm wide and distinguish vessels that are 20 μm apart deep into the brain (>8 mm). In the future, the resolution could be further improved by localizing the microbubbles directly from radio-frequency data, which could also allow the correction of aberrations from the skull[24,25].

In conventional clinical ultrasound imaging applications, resolution is inherently correlated to the ultrasonic frequency and, consequently, is inversely correlated to penetration depth. However, in uULM, resolution is related to the signal-to-noise ratio, the bandwidth of backscattered echoes and the number of array elements used in the beamforming process. This indicates that very high resolution could be reached, even deep into organs, in clinical applications. As microbubbles are clinically approved contrast agents and our acoustic parameters are well within the US Food and Drug Administration guidelines, such clinical applications could be rapidly implemented with conventional transducers. For these reasons, it is conceivable that dynamic images of the human brain vasculature could be achieved with lower frequency ultrasound (around 1 MHz) that can penetrate the skull. Ultrafast ultrasound localization could also be applied to other deep-seated organs such as liver, kidney or breast, currently imaged with ultrasound by implementing appropriate motion-correction algorithms. Such algorithms can be performed through image registration based on the cross-correlation of the radio-frequency signal acquired at high frame rates, which can detect motion at the micrometric scale[5,14,15]. The microbubble events necessary for uULM can then be motion compensated thanks to this co-registered image. Consequently, this technique will probably have an important impact on the study and diagnostics of normal biological processes or diseases such as tumour-related angiogenesis.

We demonstrate super-resolution images of rat brain microvessels with pixel sizes comparable to the size of red blood cells, indicating that vessels ten times smaller than the ultrasonic wavelength can be mapped. Since ultrafast localization imaging can be performed through the skull, non-invasive longitudinal studies may be envisioned in the future over single or multiple planes within very reasonable acquisition times in preclinical or clinical studies. ULM, by removing the diffraction-induced trade-off between resolution and penetration of ultrasound waves, emerges as the first *in vivo* technique for imaging and quantifying blood flow at microscopic resolution deep into living organs.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 12 June; accepted 30 September 2015.

1. Hell, S. W. & Wichmann, J. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Opt. Lett.* **19,** 780–782 (1994).
2. Betzig, E. *et al.* Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313,** 1642–1645 (2006).
3. Huang, B., Babcock, H. & Zhuang, X. Breaking the diffraction barrier: super-resolution imaging of cells. *Cell* **143,** 1047–1058 (2010).
4. Tanter, M. & Fink, M. Ultrafast imaging in biomedical ultrasound. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **61,** 102–119 (2014).
5. Couture, O., Tanter, M. & Fink, M. Method and device for ultrasound imaging. Patent Cooperation Treaty (PCT)/FR2011/052810 (2010).
6. Desailly, Y., Couture, O., Fink, M. & Tanter, M. Sono-activated ultrasound localization microscopy. *Appl. Phys. Lett.* **103,** 174107 (2013).
7. Couture, O. *et al.* Ultrafast imaging of ultrasound contrast agents. *Ultrasound Med. Biol.* **35,** 1908–1916 (2009).
8. Chugh, B. P. *et al.* Measurement of cerebral blood volume in mouse brain regions using micro-computed tomography. *Neuroimage* **47,** 1312–1318 (2009).
9. Huang, C.-H. *et al.* High-resolution structural and functional assessments of cerebral microvasculature using 3D Gas ΔR_2*-mMRA. *PLoS One* **8,** e78186 (2013).
10. Hong, G. *et al.* Multifunctional *in vivo* vascular imaging using near-infrared II fluorescence. *Nature Med.* **18,** 1841–1846 (2012).
11. Yao, J. *et al.* High-speed label-free functional photoacoustic microscopy of mouse brain in action. *Nature Methods* **12,** 407–410 (2015).
12. Gessner, R. C., Frederick, C. B., Foster, F. S. & Dayton, P. A. Acoustic angiography: a new imaging modality for assessing microvasculature architecture. *Int. J. Biomed. Imaging* **2013,** 936593 (2013).
13. Demene, C. *et al.* Spatiotemporal clutter filtering of ultrafast ultrasound data highly increases Doppler and fUltrasound sensitivity. *IEEE Trans. Med. Imaging* **PP**, 2271–2285 (2015).
14. Tanter, M., Bercoff, J., Sandrin, L. & Fink, M. Ultrafast compound imaging for 2-D motion vector estimation: application to transient elastography. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **49,** 1363–1374 (2002).
15. Denarie, B. *et al.* Coherent plane wave compounding for very high frame rate ultrasonography of rapidly moving targets. *IEEE Trans. Med. Imaging* **32,** 1265–1276 (2013).
16. Viessmann, O. M., Eckersley, R. J., Christensen-Jeffries, K., Tang, M. X. & Dunsby, C. Acoustic super-resolution with ultrasound and microbubbles. *Phys. Med. Biol.* **58,** 6447–6458 (2013).
17. Christensen-Jeffries, K., Browning, R. J., Tang, M.-X., Dunsby, C. & Eckersley, R. J. *In vivo* acoustic super-resolution and super-resolved velocity mapping using microbubbles. *IEEE Trans. Med. Imaging* **34,** 433–440 (2015).
18. Paxinos, G. & Watson, C. *The Rat Brain in Stereotaxic Coordinates* 6th edn (Academic, 2006).
19. Szabo, T. L. in *Diagnostic Ultrasound Imaging* (ed. Szabo, T. L.) 337–380 (Academic, 2004).
20. Mishra, A. *et al.* Imaging pericytes and capillary diameter in brain slices and isolated retinae. *Nature Protocols* **9,** 323–336 (2014).
21. Itoh, Y. & Suzuki, N. Control of brain capillary blood flow. *J. Cereb. Blood Flow Metab.* **32,** 1167–1176 (2012).
22. Kamoun, W. S. *et al.* Simultaneous measurement of RBC velocity, flux, hematocrit and shear rate in vascular networks. *Nature Methods* **7,** 655–660 (2010).
23. Goss, S. A., Frizzell, L. A. & Dunn, F. Ultrasonic absorption and attenuation in mammalian tissues. *Ultrasound Med. Biol.* **5,** 181–186 (1979).
24. Pernot, M., Montaldo, G., Tanter, M. & Fink, M. 'Ultrasonic stars' for time-reversal focusing using induced cavitation bubbles. *Appl. Phys. Lett.* **88,** 034102 (2006).
25. O'Reilly, M. A. & Hynynen, K. A super-resolution ultrasound method for brain vascular mapping. *Med. Phys.* **40,** 110701 (2013).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported principally by the Agence Nationale de la Recherche (ANR), within the project ANR MUSLI. We thank the Fondation Pierre-Gilles de Gennes for funding C.E. The laboratory was also supported by LABEX WIFI (Laboratory of Excellence ANR-10-LABX-24) within the French Program "Investments for the Future" under reference ANR-10-IDEX-0001-02 PSL*.

Author Contributions C.E., J.P., S.P., Z.L., O.C. and M.T. designed the experiments; C.E. and S.P. performed the experiments; C.E., J.P. and O.C. analysed the data; Y.D., O.C. and M.T. performed the theoretical analysis. All the authors discussed the results and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.T. (mickael.tanter@espci.fr).

METHODS

Theoretical resolution limit. The given theoretical resolution limit corresponds to the position error of the localization process[6]. This PSF deconvolution for single isolated spots is inherently limited by the number of channels used in receive processing and the timing resolution of the acquisition system. The latter is limited mainly by the sampling frequency of the echoes before beamforming. An approximate value for the theoretical resolution limit in the axial dimension can be obtained by propagating the sampling error in a time-of-flight model, which yields:

$$\sigma_{z0} \approx c\sigma_\tau/(2n^{1/2})$$

where σ_{z0} is the localization error in the axial dimension, c is the sound speed, σ_τ is the timing resolution of the system and n is the number of channels used in receive processing. Note that the lower limit of the timing resolution is linked to the Cramer–Rao lower bound (CRLB), which describes the minimum obtainable estimation error variance when using an unbiased estimator. The derivation of the CRLB was given by Walker and Trahey for ultrasound[26]. For tasks related to ultrasonic displacement estimation. The standard deviation σ_τ of arrival time estimates compared to the theoretical one is described by the relation:

$$\sigma_\tau = \sigma\left(\hat{\tau}_0 - \tau_0\right) \geq \sqrt{\frac{3}{2f_0^3\pi^2 T\left(B^3 + 12B\right)}\left(\frac{1}{\rho^2}\left(1+\frac{1}{\mathrm{SNR}^2}\right)^2 - 1\right)}$$

where f_0 is the transmit pulse centre frequency, B is the pulse bandwidth, T is the kernel size for the time delay estimation, ρ is the normalized correlation between signals (that is, the correlation between the experimental signal and the reference signal used for the PSF decorrelation), and SNR is the signal-to-noise ratio of receive signals.

For the lateral resolution, the size of the aperture must also be taken into account as in any classical imaging modality:

$$\sigma_{x0} \approx 2 \times 3^{1/2} c\sigma_\tau f/(Dn^{1/2})$$

where σ_{x0} is the localization error in the lateral dimension, f is the focal length and D is the length of the transducer array (which is the imaging aperture here). Following these theoretical models, it is predicted that the 15 MHz array used in this study could attain a maximum resolution (full-width at half-maximum) of 2.5 μm in the axial direction and 5 μm in the lateral direction at 1 cm depth. In humans, lower frequencies are exploited to attain 10 cm penetration. With the same theoretical model, we can predict a 6 μm isotropic resolution with a current transducer matrix (32 × 32 elements, 300 μm spatial pitch, 2.5 MHz frequency, 70% frequency bandwidth, $\rho > 0.9$, 12 dB SNR at 5 cm depth).

Animals. All experiments were performed in agreement with the European Community Council Directive of 22 September 2010 (010/63/UE) and the local ethics committee (Comité d'éthique en matière d'expérimentation animale no. 59, C2EA–59, 'Paris Centre et Sud'). Accordingly, the number of animals in our study was kept to the necessary minimum. Experiments were performed on $n = 3$ male Sprague–Dawley rats (Janvier Labs), weighing 200–225 g at the beginning of the experiments. Animals arrived in the laboratory 1 week before the beginning of the experiment, and were housed three per cage. They were kept at a constant temperature of 22 °C, with a 12 h alternating light/dark cycle (light 7 a.m. to 7 p.m.). Food and water were available *ad libitum*.

Preparation of the thinned-skull imaging windows. The skull of the rats was thinned to 75–100 μm over an area of approximately 0.6 cm × 0.9 cm. The thinned window suits the dimension of the ultrasound linear array (0.08 mm per element; 128 elements = 10.24 mm width). The surgical procedure was performed 1–2 days before imaging under anaesthesia using intraperitoneal injections of medetomidine (Domitor; 0.3 mg kg^{-1}) and ketamine (Imalgène; 40 mg kg^{-1}). The head of the animal was placed in a stereotaxic frame and the skull bone was drilled (Foredom) at low speed with a micro drill steel burr (Fine Science Tools, catalogue no. 19007-07). To prevent swelling, or oedema of the cerebral cortex, the skull was frequently cooled with saline and an airstream during the thinning procedure as described previously[27]. The thinned window was protected by a small (1 cm × 1 cm) plastic cover, and the skin was sutured using 5.0 non-absorbable Ethicon thread. Preliminary experiments showed that this method enabled good quality ultrasound imaging results within 24 h to 3 days after the preparation, as the bone tends to re-grow.

Preparation of ultrasound contrast-agent microbubbles. To reconstruct the vascular microstructure of the rat brain, 1–5 μm perfluorocarbon-filled microbubbles (Bracco) were dissolved with 0.9% NaCl to yield an initial concentration of 2×10^8 microbubbles per ml. This concentration corresponds to approximatively 500,000 bubbles per ml of blood per injection, which corresponds to the maximal dose injected in clinical practice for superficial contrast-enhanced ultrasound[28]. Ultrafast ultrasound localization microscopy was performed in the brain by injecting a maximum of 18 bolus injections (corresponding to 2.7 ml of the initial suspension) through the catheterized jugular vein. The coronal ultrafast acquisitions of the brain were performed every 15 min to guarantee that the injected boluses had been cleared out.

Ultrafast ultrasound imaging sequence. Ultrasound imaging was performed using ultrafast Doppler imaging based on compounded plane-wave ultrasound transmissions[29,30]. The hardware of the ultrasound scanner was not modified. Ultrafast sequences were initiated and processed through software-based sequence encoding and data were imported through a PCI-Xpress fast bus for GPU-based post-treatment.

Owing to its high spatiotemporal resolution (1 ms, 100 μm) (ref. 31), this technique can measure small haemodynamic changes related to the neurovascular coupling. Real-time B-mode imaging was used to control the placement of the probe on the field of view. In detail, we developed a plane-wave compounded ultrafast imaging sequence (three tilted plane waves, −3°, 0° and 3°, pulse-repetition frequency PRF = 1,500 Hz) to perform a scan of the entire brain and have a detailed overview of its microvasculature over different coronal imaging planes at a high frame rate (500 Hz). Our ultrasonic probe is a custom-built array with 160 elements and a central frequency of 20.3 MHz (pitch = 0.08 mm, elevation focus = 10 mm). Its 15.4 MHz bandwidth allowed the use of this probe at a frequency of 15 MHz. The signal from the 16 elements on either side is discarded as it is mounted on a fully programmable ultrasound clinical scanner with 256 channels in transmission and 128 parallel channels in reception. Data are transferred using a 16×, 6 Gb s^{-1} PCI express bus and processed using a 12-core 3 GHz Xeon processor, NVidia Quadro K5000 Graphical Processing Unit with a bus at 173 Gb s^{-1}, providing 2.1 teraflops. Such software-based architecture enables programming of custom transmit/receive sequences where the frame rate of each acquisition can reach more than 20 KHz. The linear array was coronally fixed at the anterior–posterior coordinates of Bregma −0.5 mm and coronally translated for 500 μm with a motor to scan and retrieve the vasculature of the whole brain along 2 cm. Each pressure transmit pulse consists of 6 cycles (2 μs duration at 15 MHz) at a 1.5 MPa peak rarefaction acoustic pressure (mechanical index = 0.4). These pressure amplitudes are chosen to reduce the ultrasound-induced disruption of microbubbles and to allow the tracking of these agents over several images.

Boluses of 150 μl microbubbles were injected at the beginning of each ultrafast acquisition. Once the scan was completed, we fixed the probe above the Bregma −1.0 mm to continuously insonify for 150 s the rat cortex (3.5 mm depth). Ten minutes of acquisition were required per each coronal plane of the whole-brain scan (11.6 mm depth). In this latter case, we injected two 150 μl boluses of contrast agents (at the beginning of the ultrafast acquisition and in the middle, 5 min) to avoid a drop in the microbubble concentration due to the dynamic of the boluses. The backscattered echoes were recorded, beamformed with λ-line spacing and coherently added to produce an echographic image at each transmission. Successive raw images corresponding to three different transmission angles at 1,500 Hz PRF are then coherently added to produce one higher-contrast ultrasonic image for each set of tilted angles at a 500 Hz frame rate.

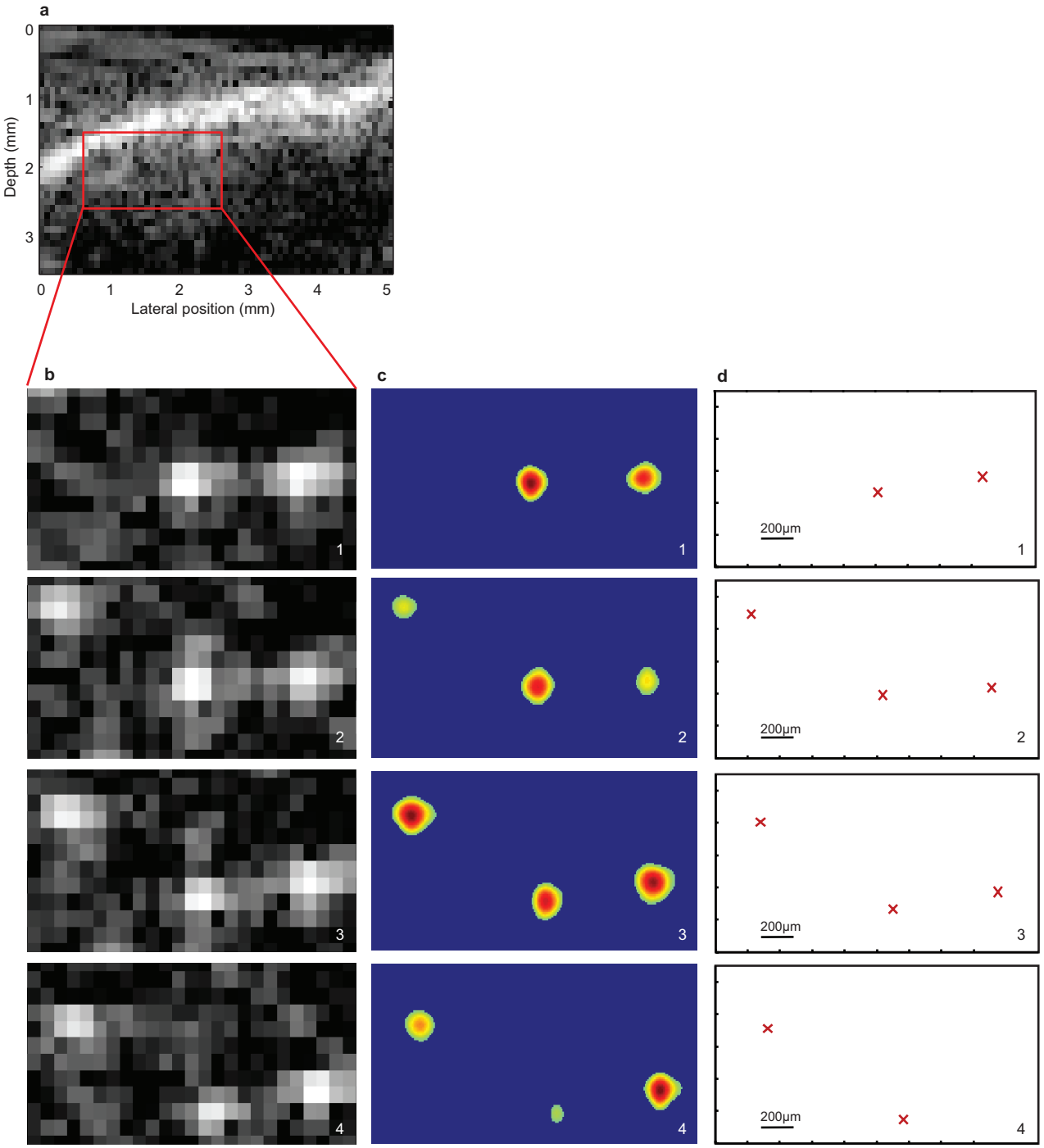
Data treatment for bubble localization. High-pass spatiotemporal filtering was implemented on the stack of the ultrafast images to discriminate the high temporal components, belonging to the blood signal, from the slow-moving tissue. Next, the stack of filtered ultrafast acquisition was rescaled via interpolation, yielding super-resolved output images with a pixel size of 10 μm × 8 μm. Since the bubbles are much smaller than the wavelength (1–3 μm versus 100 μm) and can be individually separated in space and time, they appear as the PSF of the ultrasound system. This PSF is well behaved with respect to the theory of acoustic diffraction because human and animal soft tissues can be considered homogeneous for acoustic properties at first-order approximation[32].

Thereafter, we computed a Gaussian low-pass spatial filter and extracted a two-dimensional PSF for deconvolution of the rescaled ultrafast acquisitions. Hence, each individual bubble was localized, across all frames in the axial position and in depth, with a Gaussian two-dimensional profile whose summit represents the centroid of each separable source (Extended Data Fig. 1). Only 50% of the maximum of the full-width at half-maximum was kept to reconstruct the density maps of the bubbles; such thresholding helped cancel unwanted noisy signals. Additionally, to avoid any artefact corresponding to independent neighbouring bubble events, only bubbles that could be followed for at least 2 ms were included. Eventually the bubbles were counted and grouped according to their closeness. Almost 1.2 million bubbles were counted in the rat cortex within 74,800 frames. Supplementary Video 1 shows the reconstruction of each vessel through the passage of individual microbubbles.

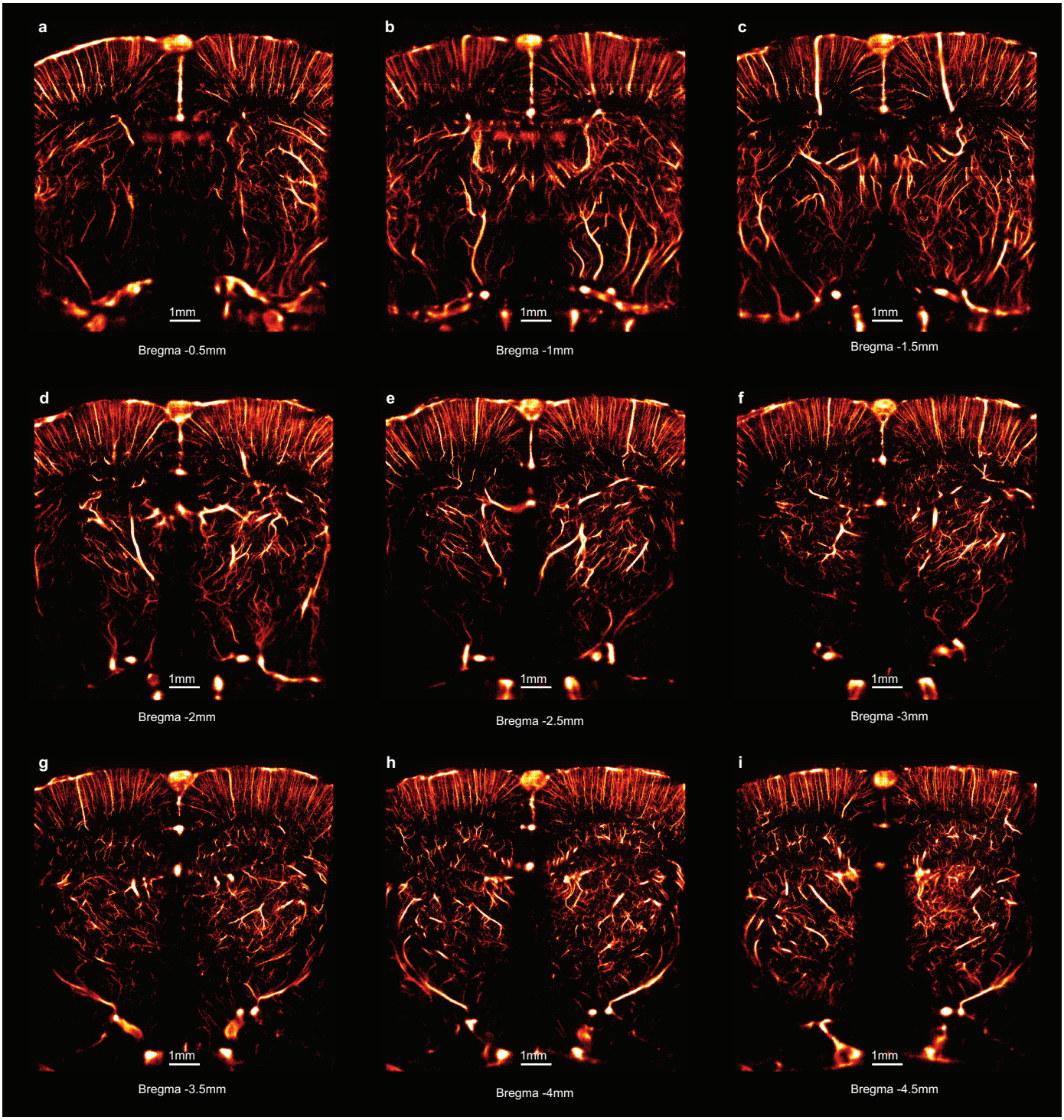
A displacement vector was drawn between these positions, enabling the evaluation of the instantaneous in-plane velocities of the bubbles, computed as the rate

of displacement from one frame to the next frame divided by the time interval. Only tracks composed by more than 5 frames (10 ms) were considered to evaluate the velocities. Coloured velocity maps were constructed using the bubble paths associated with their in-plane velocities (Fig. 2d). More specifically, blue corresponds to the velocities towards the top and the red refers to the in-plane velocities towards the bottom. Taken separately but treated equally, the velocity maps were exploited to retrieve the velocity profiles of each downstream and upstream micro-vessel. In Fig. 2e, we selected two representative vessels: (4) and (5), whose velocities were oriented towards the bottom and towards the top, respectively. We evaluated the number of bubbles in a fixed-resolution cell (Δxz), across the sections of the two chosen vessels, and extracted 50% of the fastest bubbles. Then, we measured the mean $\pm$ standard deviation of each thresholded in-plane velocity vector and performed an unpaired Student's t-test. When Δxz was chosen between 8 μm and 12 μm, the quantification of the velocity distribution for each resolution cell gave a result that was statistically different from the adjacent one ($P < 0.05$). Finally, uULM was performed to reconstruct the vascular network and quantify the velocity maps in the whole brain. In Extended Data Fig. 2, we show how the microvasculature of the brain was retrieved with high resolution in depth (11.6 mm) along different coronal imaging planes (from Bregma -0.5 mm to Bregma -4.5 mm). Each of these ultrasound acquisitions was detached in three panels of 4 mm depth to properly filter out the thinned skull bone. Supplementary Video 2 shows the various coronal slices taken during the experiments. It should be noted that the same filter was applied to reconstruct the vasculature of the cortex in Figs 2a and 3a, and Extended Data Fig. 2a–i. The in-plane velocity maps in Extended Data Fig. 3 were attained with the same data treatment as Figs 2d and 3b. They enable the quantification of velocity distributions in depth in the whole brain, corresponding to the coronal imaging plans in Extended Data Fig. 3.

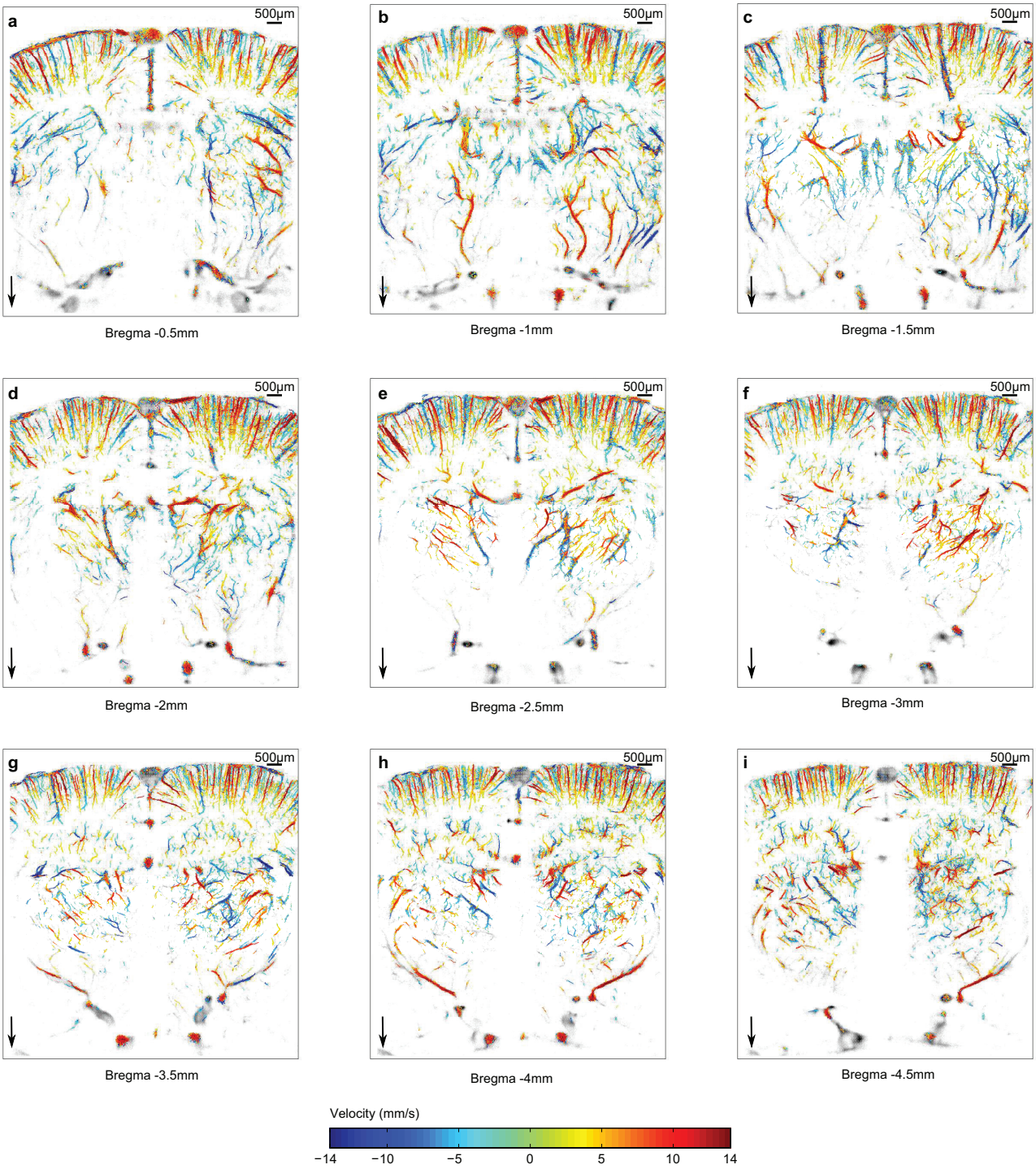
26. Walker, W. F. & Trahey, G. E. A fundamental limit on delay estimation using partially correlated speckle signals. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **42,** 301–308 (1995).
27. Osmanski, B.-F., Pezet, S., Ricobaraza, A., Lenkei, Z. & Tanter, M. Functional ultrasound imaging of intrinsic connectivity in the living rat brain with high spatiotemporal resolution. *Nature Commun.* **5,** 5023 (2014).
28. Dietrich, C. F. *et al.* An EFSUMB introduction into Dynamic Contrast-Enhanced Ultrasound (DCE-US) for quantification of tumor perfusion. *Ultraschall Med.* **33,** 344–351 (2012).
29. Bercoff, J. *et al.* Ultrafast compound Doppler imaging: providing full blood flow characterization. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **58,** 134–147 (2011).
30. Errico, C., Osmanski, B.-F., Pezet, S., Couture, O., Lenkei, Z. & Tanter, M. Transcranial functional ultrasound imaging of the brain using microbubble-enhanced ultrasensitive Doppler. *NeuroImage* **124,** 752–761 (2015).
31. Macé, E. *et al.* Functional ultrasound imaging of the brain. *Nature Methods* **8,** 662–664 (2011).
32. Duck, F. A. *Physical Properties of Tissues: A Comprehensive Reference Book* (Academic, 2013).



Extended Data Figure 1 | Schema of the temporal and spatial localization of unique sources. a, Stack of B-mode images. The region of interest corresponds to a region of 2 mm × 1.1 mm within the cortex. **b**, Spatiotemporal filtering of the B-mode images shows the presence of decorrelating microbubbles in each frame (1–4). **c**, The four representative frames are separated by 44 ms (1–4). **d**, Computed two-dimensional PSF of the rescaled and filtered ultrafast acquisitions. These echoes are then interpolated and the Cartesian coordinates of their centre is obtained (1–4). The summit of each two-dimensional Gaussian profile identifies the centroid of each separable source.



Extended Data Figure 2 | uULM coronal scan (anterior–posterior) of the entire rat brain through a thinned skull window. a–i, The ultrasound probe was driven by a micro-step motor to perform uULM on different imaging planes separated by 500 µm. We reconstructed the vascularization of the rat brain at the following coordinates: Bregma −0.5 mm (**a**), −1 mm (**b**), −1.5 mm (**c**), −2 mm (**d**), −2.5 mm (**e**), −3 mm (**f**), −3.5 mm (**g**), −4 mm (**h**), −4.5 mm (**i**).



Extended Data Figure 3 | Anterior–posterior scan of in-plane velocity maps of the rat forebrain through a thinned skull window. a–i, Velocity maps for the different coronal planes presented in Extended Data Fig. 2.

LETTER

doi:10.1038/nature15734

Extra adsorption and adsorbate superlattice formation in metal–organic frameworks

Hae Sung Cho[1]*, Hexiang Deng[2,3]*, Keiichi Miyasaka[1]*, Zhiyue Dong[2], Minhyung Cho[1], Alexander V. Neimark[4], Jeung Ku Kang[1], Omar M. Yaghi[1,5,6] & Osamu Terasaki[1,7]

Metal-organic frameworks (MOFs) have a high internal surface area and widely tunable composition[1,2], which make them useful for applications involving adsorption, such as hydrogen, methane or carbon dioxide storage[3–9]. The selectivity and uptake capacity of the adsorption process are determined by interactions involving the adsorbates and their porous host materials. But, although the interactions of adsorbate molecules with the internal MOF surface[10–17] and also amongst themselves within individual pores[18–22] have been extensively studied, adsorbate–adsorbate interactions across pore walls have not been explored. Here we show that local strain in the MOF, induced by pore filling, can give rise to collective and long-range adsorbate–adsorbate interactions and the formation of adsorbate superlattices that extend beyond an original MOF unit cell. Specifically, we use *in situ* small-angle X-ray scattering to track and map the distribution and ordering of adsorbate molecules in five members of the mesoporous MOF-74 series along entire adsorption–desorption isotherms. We find in all cases that the capillary condensation that fills the pores gives rise to the formation of 'extra adsorption domains'—that is, domains spanning several neighbouring pores, which have a higher adsorbate density than non-domain pores. In the case of one MOF, IRMOF-74-V-hex, these domains form a superlattice structure that is difficult to reconcile with the prevailing view of pore-filling as a stochastic process. The visualization of the adsorption process provided by our data, with clear evidence for initial adsorbate aggregation in distinct domains and ordering before an even distribution is finally reached, should help to improve our understanding of this process and may thereby improve our ability to exploit it practically.

Figure 1 shows the three distinct types of interaction in which adsorbates in MOFs can engage: adsorbate molecules can interact with the material's internal surface (regime A); adsorbates can interact among themselves within the confines of a pore (regime B); and adsorbates can interact with each other across pores mediated by the material framework (regime C). Studying the collective adsorbate behaviour in regimes B and C requires porous MOF crystals, with pores that are large enough to enable the organization and behaviour of confined adsorbates to be observed, and with pore walls that are atomically thin and well-defined so as to allow observation of any local perturbations resulting from adsorption. In such systems, we can then use *in situ* small-angle X-ray scattering (SAXS) to detect long-range ordering of adsorbates in multiple pores at precisely controlled temperatures and pressures.

We chose the five mesoporous MOFs with isoreticular structure (IRMOF-74-III, IRMOF-74-IV, IRMOF-74-V, IRMOF-74-V-hex and IRMOF-74-VII) that are based on the crystalline IRMOF-74 structure[8,23–25]. The robustness of the IRMOF-74 honeycomb-like structure (in projection) is imparted by one-dimensional, rod-shaped magnesium oxide units that run along the pore direction and are held together by organic linkers (Fig. 2a). This rigid oxide unit allows for structural refinements in two dimensions, by keeping constant the structure along the *c* axis of the original MOF structure[8] (Fig. 2b). Thus, we apply the projected symmetry of the two-dimensional space groups (plane groups) *p*3 or *p*6 for the unit cell (Fig. 2b, green parallelogram). We therefore need only two variables, *h* and *k*, to specify the reflections with the *h* and *k* indices for the refinement. This allows us to focus on the adsorption region, and stops us from having to deal unnecessarily with the more complicated original symmetry $R3$ in IRMOF-74-IV, IRMOF-74-V and IRMOF-74-VII, or $R\bar{3}$ in IRMOF-74-III and IRMOF-74-V-hex (Fig. 2b, red parallelogram).

All of these MOFs exhibit open porosity and have mesopores with sizes of 22 Å, 28 Å, 35 Å and 49 Å (for IRMOF-74-III, IRMOF-74-IV, IRMOF-74-V and IRMOF-74-VII, respectively). IRMOF-74-V-hex, having a pore size of 34 Å, was constructed with a linker functionalized

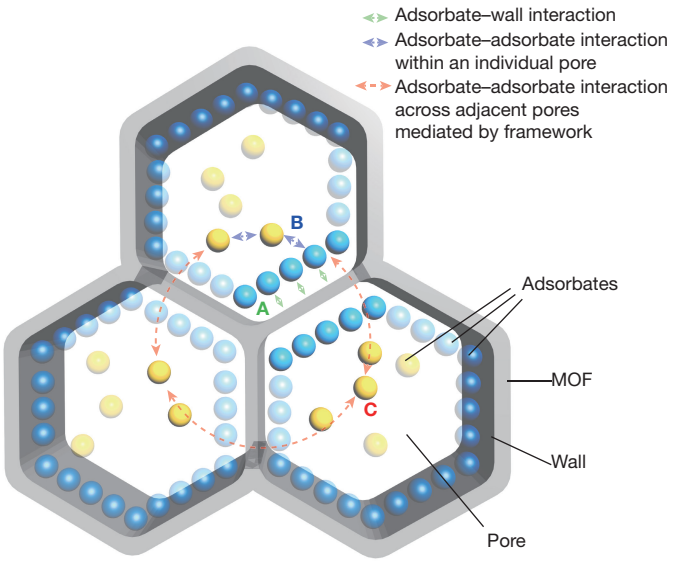


Figure 1 | Three adsorbate-interaction regimes in mesoporous MOFs. In regime A, adsorbed molecules interact (green arrows) with pore walls. In regime B, adsorbates interact amongst each other (blue arrows) within a pore. These two types of interaction and the corresponding regimes have been well studied. Regime C, however, has not been explored; here, adsorbates interact with each other (red arrows) across pore walls, in a way that is mediated by the framework. Light blue, molecules adsorbed onto the internal pore surface; yellow, molecules in the centre of the pores.

[1]Graduate School of Energy, Environment, Water and Sustainability, WCU/BK21Plus, KAIST, Daejeon 305-701, South Korea. [2]Key Laboratory of Biomedical Polymers-Ministry of Education, College of Chemistry and Molecular Sciences, Wuhan University, Luojiashan, Wuhan 430072, China. [3]The Institute for Advanced Studies, Wuhan University, Wuhan 430072, China. [4]Department of Chemical and Biochemical Engineering, Rutgers University, 98 Brett Road, Piscataway, New Jersey 08854, USA. [5]Department of Chemistry, University of California, Materials Sciences Division at Lawrence Berkeley National Laboratory, and Kavli Energy NanoSciences Institute, Berkeley, California 94720, USA. [6]King Fahd University of Petroleum and Minerals, Dhahran 34464, Saudi Arabia. [7]Department of Materials and Environmental Chemistry, Berzelii Centre EXSELENT on Porous Materials, Stockholm University, Stockholm SE-10691, Sweden.
*These authors contributed equally to this work.

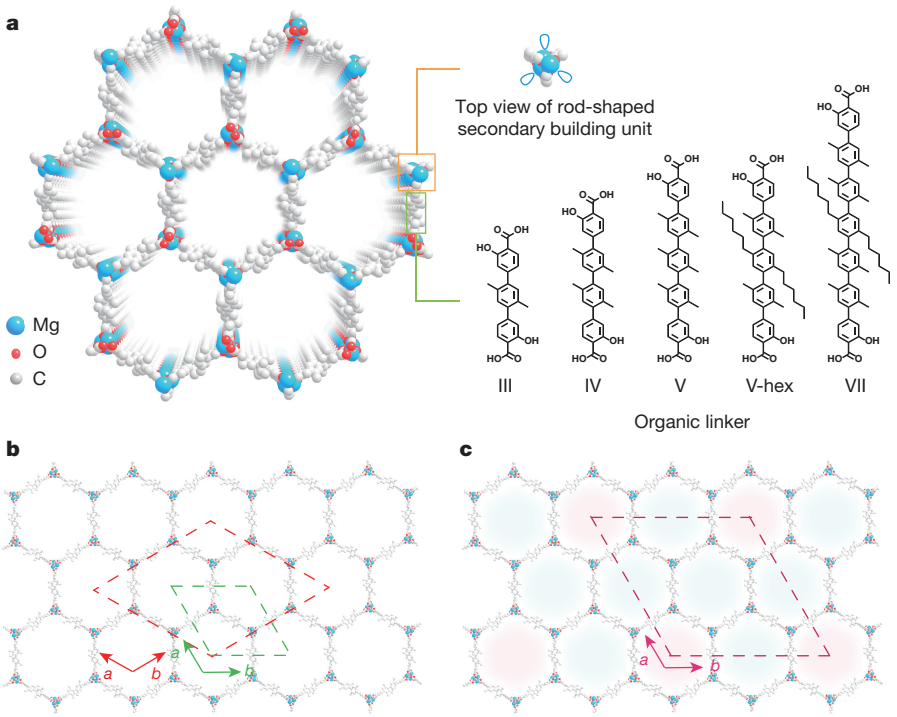


Figure 2 | Structure of the IRMOF-74 series in three and two dimensions. a, The honeycomb-like structure (in projection) of MOFs of the IRMOF-74 series is imparted by one-dimensional, rod-shaped magnesium oxide secondary building units, held together by organic linkers (III, IV, V, V-hex or VII). **b**, Green dashes show the two-dimensional unit cell, corresponding to plane groups *p*3 and *p*6, that we applied for structural refinement. Red dashes shown the $\sqrt{3} \times \sqrt{3}$ unit cell, which corresponds to the projection of the original space groups *R*3 and $R\bar{3}$, used to reveal adsorbate distribution. **c**, The 2×2 superlattice cell (purple dashes) and the pores in violet contain a larger number of adsorbates, compared with the surrounding pores in blue.

with a hexyl chain (V-hex). The atomically thin walls of the pores in these MOFs and their large pore sizes are key factors in their suitability for examining the collective behaviour of the adsorbates within and across the pores (regimes B and C, Fig. 1).

In contrast to other *in situ* adsorption studies, performed using a synchrotron beamline[16,17,26], we used a laboratory-designed SAXS set-up operating in transmission mode with a rotating anode X-ray source, a graded confocal optic, and a Kratky block system to create a monochromatic beam focusing on the detector. Incorporation of an adsorption apparatus in the SAXS system enables measurement of both X-ray-diffraction profiles and gas-adsorption isotherms from the same sample at a precisely controlled temperature and adsorbate pressure (see Supplementary Information, section 1). We illustrate SAXS-based adsorption tracking for argon uptake by IRMOF-74-V-hex, for which the adsorption process can be divided into stages 1 to 5, taking place within the pressure ranges 0 to 0.5 kPa, 0.5 to 27 kPa, 27 to 33 kPa, 33 to 50 kPa and 50 to 100 kPa, respectively (Fig. 3a). Although the shape of the isotherm is similar to that of a type IV isotherm (as classified by the International Union of Pure and Applied Chemistry), typical for mesoporous materials, the distinct slopes seen in stages 4 and 5 point to two major differences. To understand the origin of slopes, we measured SAXS profiles along the entire adsorption curve, among which 11 different gas pressures (Fig. 3a) were selected to represent the different stages of the argon adsorption process (Fig. 3b).

The electron distribution of argon atoms introduced into the pores was obtained from 54 (7 independent) reflections, using difference Fourier analysis of the measured intensity profile of the argon-filled MOF and the calculated intensity profile of the corresponding activated MOF without argon (Fig. 3d; see also Supplementary Figs 4 and 5 and Tables 1–4). (Note that, although we cannot determine the argon distribution with atomic resolution owing to limitations imposed by the maximum q range (the largest angle that can be detected; $4\pi\sin\theta/\lambda$) of the SAXS instrument, the resolution is sufficient to map the electron-distribution trend within the large pores; an exception is the centre point, which might be slightly affected by the termination effect in the Fourier synthesis.) The adsorbate electron-distribution maps (Fig. 3d) reveal that, as expected and in agreement with previous findings[10], argon interacts strongly with the open metal sites of the magnesium oxide units in stages 1 and 2. The electron-density map of argon at 27 kPa (Fig. 3d) and corresponding electron density distribution profile (Fig. 3c) show two to three cylindrical layers of argon atoms adsorbed onto the walls (regime A in Fig. 1). This is followed by argon condensation in the pores, which commences at stage 3 and is accompanied by a steep increase in gas uptake (Fig. 3a, d). At roughly midway through stage 3, the corresponding $hk = 10$ reflection intensity decreases sharply (Fig. 3b; Supplementary Fig. 10), while a new broad peak at $q = 0.10$ Å^{-1} emerges (marked by grey dashes in Fig. 4a)—evidence of collective adsorbate–adsorbate interactions). Although the pores are not yet completely filled, as indicated by the smeared-out electron density in the centre region of the pore (Fig. 3d), the emergence of this broad peak unambiguously represents an important point (termed the aggregation point) in the initiation of formation of extra adsorption domains, whereby adsorbate atoms gather in certain pore regions in higher numbers than the average.

The intensity of this broad peak reaches a maximum as stage 3 turns to stage 4 (33 kPa), and then decreases gradually to eventually disappear at the end of stage 4 (Fig. 4a). Furthermore, the density of argon in the centre region increases more than it does around the pore walls (Fig. 3d, stage 4) during both the appearance and the disappearance of the peak. The correspondence between the appearance/disappearance of this peak with the characteristics of the pore-filling process indicates that this unusual phenomenon originates from the complex collective behaviour of argon: argon atoms are not equally distributed throughout the available pores during stages 3 and 4 of the condensation process, but instead exhibit density fluctuations that result in extra adsorption domains with a higher-than-average argon concentration spanning several contiguous pores.

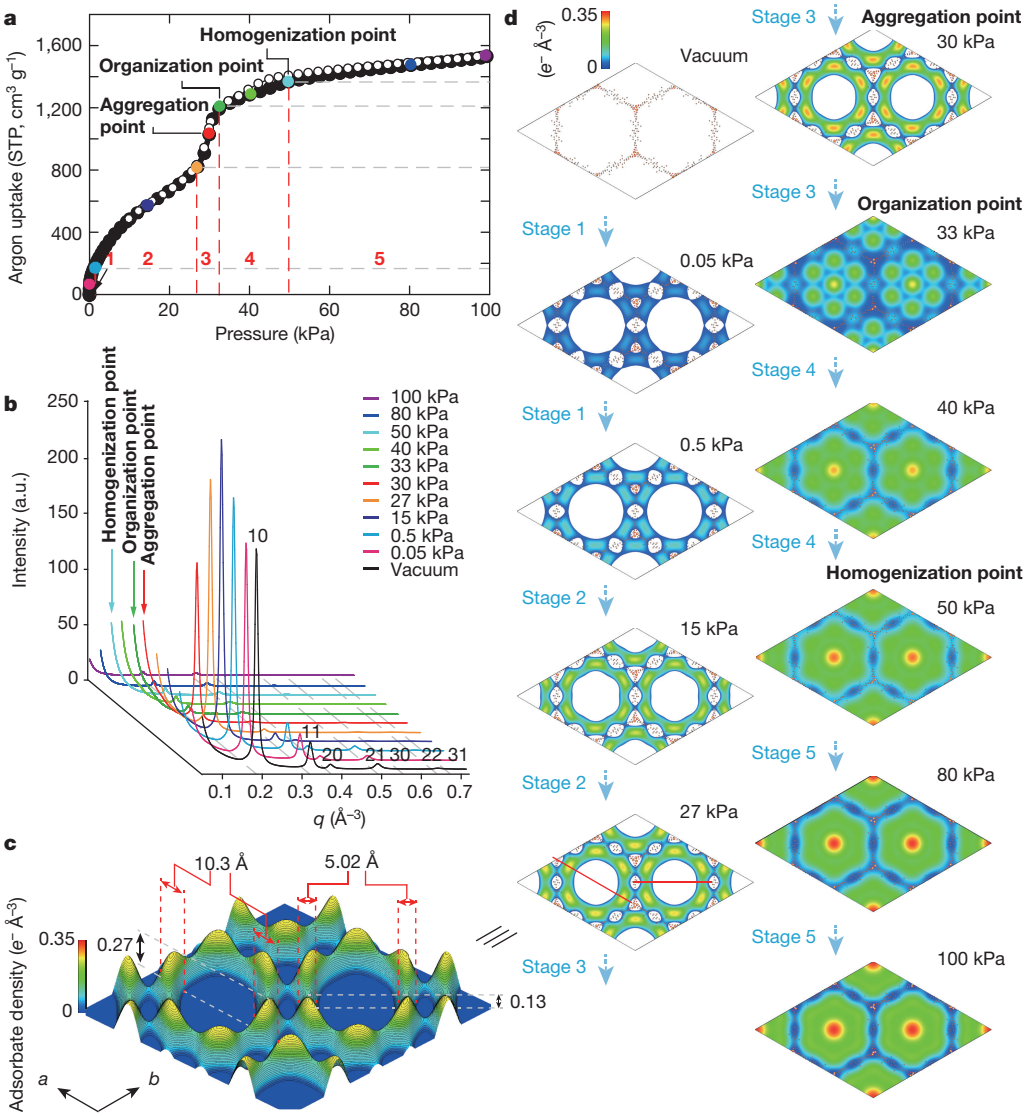


Figure 3 | Mapping of argon distribution in IRMOF-74-V-hex. a, Argon uptake by IRMOF-74-V-hex at different gas pressures. The isotherm shows five stages (1 to 5), with distinct slopes. Three points (red, dark green and light blue) are highlighted for the start, end/start and end of two events unobserved in type IV isotherms. **b**, SAXS scattering profiles measured along the entire adsorption process at 11 different gas pressures, covering the different stages of argon adsorption. The patterns are overlaid in linear scale, with colours corresponding to the points in the isotherm. **c**, Three-dimensional contour map of the electron-density profile of argon at 27 kPa. The mesopores are covered by argon at this point. **d**, Projected argon distribution in two dimensions from the three-dimensional contour maps. Each two-dimensional map reveals the argon distribution within the MOF structure at a certain pressure. The red lines at 27 kPa indicate the argon profile projection in two directions (metal site to metal site, and wall to wall).

Further evidence for the formation of these domains comes from the adsorption profile slope in stage 4, which differs strongly from that in stage 3. The formation of the domains causes unit-cell contraction and associated broadening of the diffraction intensity profiles ($hk = 10$, 11, 20, 21, 30, 22 and 31) during stage 3, and expansion of the unit cell and sharpening of the associated profiles during stage 4 (Fig. 4b–d). These effects are correlated with changes in the local strain of the MOF backbone, which results in the emergence of an additional stage in the overall adsorption process in IRMOF-74, as indicated by the full-width half-maxima (FWHM) of the SAXS profile peaks, and by the unit-cell parameters (contraction versus expansion) adopting during stage 3 maximum and minimum values, respectively (Fig. 4c, d). In stage 4, the strain induced by the adsorption heterogeneity starts to smear out and the FWHM decreases as more argon atoms enter the pores and move towards a more homogenized arrangement, leading to a different slope for gas uptake[27]. We note that the changes in the FWHM and unit-cell parameters of IRMOF-74 during and after mesopore condensation resemble those accompanying gas adsorption in MCM-41 (a typical mesoporous silica with relatively thick pore walls[27]), but the magnitude of the changes in our system is much larger than that observed for MCM-41. This indicates that the adsorbates stress the IRMOF-74 framework, with its thinner MOF walls, more than they stress the more-sturdy MCM-41 framework; this also explains why a broad peak at low q range and a unique slope at stage 4 could not be observed during and after mesopore condensation for MCM-41 (Supplementary Fig. 12).

The fate of the extra adsorption domains in IRMOF-74-V-hex can be gleaned from the abrupt appearance of superlattice reflections in the SAXS patterns (reflections at $q = 0.25\ \text{Å}^{-1}$ (marked by grey dashes) in Fig. 4a, and at $q = 0.42\ \text{Å}^{-1}$ in Supplementary Fig. 40) at the start of stage 4. The intensity of the reflections decreases as the pressure increases from 33 kPa, and becomes zero at 50 kPa (Fig. 4b), accompanied by a decrease in the FWHM of all the profile reflections. We infer from these observations that extra adsorption domains form, and that the contrast between these domains and the surrounding domains increases

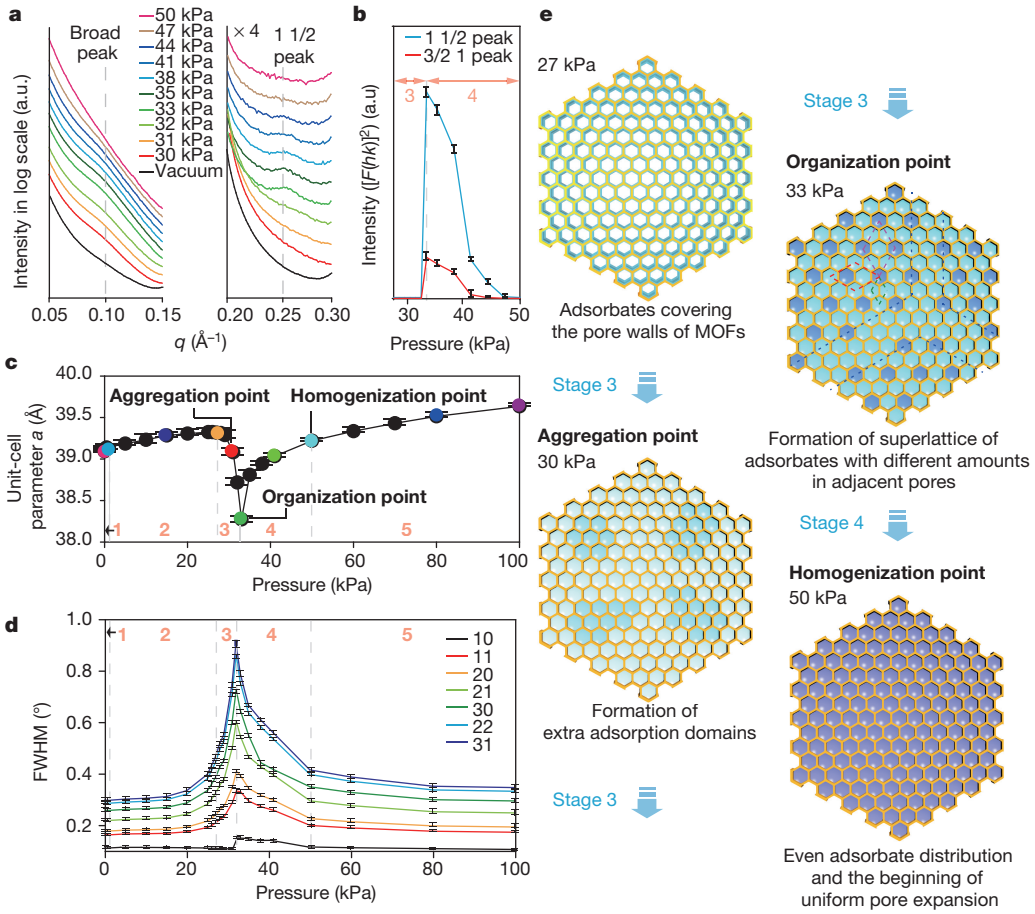


Figure 4 | Extra adsorption domains and argon adsorbate superlattice in IRMOF-74-V-hex. a, The appearance and disappearance of the broad peak (at $q = 0.093$ Å^{-1}) and the superlattice peak (at $q = 0.25$ Å^{-1}) in the SAXS patterns during the absorption process, with intensity magnified by four in the right-hand image. **b**, Intensity changes of $1\frac{1}{2}$ and $\frac{3}{2}1$ superlattice reflections. **c**, Tracking of the unit-cell parameter change of IRMOF-74-V-hex during the adsorption process. **d**, Tracking of the corresponding FWHM of reflections in the SAXS profile. **e**, Illustration of the extra adsorption domains (aggregation point, 30 kPa) and superlattice (organization point, 33 kPa) formed as a result of argon being distributed unevenly among adjacent mesopores. Green, red and purple dashes indicate the original MOF, $\sqrt{3} \times \sqrt{3}$, and 2×2 unit cell, respectively. The size of the argon superlattice domain (dark blue dashes at 33 kPa) is about 400 Å.

during stage 3 of the adsorption process (Supplementary Fig. 11); moreover, we conclude that once the point of maximum contrast (termed the organization point) has been reached, the domains commence to form the adsorbate superlattice in stage 4. It is the superlattice formation that relieves local strain—increasingly so as the contrast lessens and as more adsorbates fill the pores towards uniform distribution (Supplementary Fig. 11).

The precise structure of the superlattice is determined from the positions of the reflections at $q = 0.25$ Å^{-1} and $q = 0.42$ Å^{-1}, mentioned above, indexed by $hk = 1\frac{1}{2}$ and $\frac{3}{2}1$ and corresponding to an adsorbate superlattice structure with a 2×2 unit cell (Fig. 2c, purple parallelogram, and Fig. 4e). Although this structure also gives a reflection of $\frac{1}{2}0$ ($q = 0.093$ Å^{-1}) that overlaps with the broad peak, the position and intensity of the two observed peaks rule out the possibility of a $\sqrt{3} \times \sqrt{3}$ superlattice (Fig. 2b, red parallelogram, and Fig. 4e) that might form through modulation of the MOF structure. In such a case, the corresponding ordered reflection would have appeared at $q = 0.29$ Å^{-1} for $hk = \frac{4}{3}\frac{1}{3}$; however, these were absent in the SAXS patterns. Note also that the line-widths of the adsorbate superlattice reflections are much larger than those of fundamental reflections, further suggesting that the origin of these extra peaks is not associated with the framework lattice. Detailed analysis of the FWHM revealed that the size of the superlattice domains is about 400 Å.

Upon further increases in the argon pressure, the extra adsorption domains and superlattice reflections disappear at the end of stage 4 (Fig. 3a). During the next stage (stage 5), the adsorption isotherm shows a new slope, and the electron density in the centre region of the pores gradually increases (Fig. 3d, stage 5) and leads to a slight unit-cell expansion (Fig. 4c) to accommodate more incoming argon atoms in a uniform manner among different pores. This changeover point in the isotherm (termed the homogenization point) marks the initiation of uniform pore expansion: the adsorbate superlattice disappears and homogenization of the adsorbate density takes place without involvement of the long-range adsorbate–adsorbate interactions that are mediated by local strain in the MOF framework. In terms of the amount of argon uptake, stages 4 and 5 account for up to 22% of the total uptake in IRMOF-74-V-hex.

An overview of how different SAXS characteristics document the different stages in the overall adsorption process is provided in Extended Data Fig. 1. The desorption process of argon in IRMOF-74-V-hex—which was also carefully studied (Fig. 3a) and compared with the adsorption process in detail (Supplementary Figs 7–9, Supplementary Tables 5 and 6, and Supplementary Video)—involves the same stages as those seen during adsorption.

The broad peak that is seen at low q values was observed in all IRMOFs for all three adsorbates studied (argon, nitrogen and carbon dioxide) (Supplementary Fig. 14). During stage 3, this peak was observed in the SAXS intensity profiles at $q = 0.12$ Å^{-1} and $q = 0.094$ Å^{-1} for IRMOF-74-IV and IRMOF-74-V, respectively. From the distance distribution function derived from the SAXS data in the q range of 0.016 Å^{-1} to 0.18 Å^{-1} for IRMOF-74-IV, and 0.016 Å^{-1} to 0.16 Å^{-1} for IRMOF-74-V and IRMOF-74-V-hex, the

maximum size of individual extra adsorption domains is calculated to be approximately 60 Å for IRMOF-74-IV, and 70 Å for IRMOF-74-V and IRMOF-74-V-hex. Although extra adsorption domains were seen in all IRMOF-74 compounds during the pore-filling process, the intensity of the additional reflections that are attributed to superlattice formation was negligible in the case of IRMOF-74-IV and IRMOF-74-V. The hexyl chains of IRMOF-74-V-hex thus seem to be important in superlattice formation, although pore size will also be relevant (as superlattices were not detected in IRMOF-74-VII, where hexyl chains are present but within the confines of larger pores).

The changes in the SAXS profiles seen during adsorption and desorption of all three adsorbates follow similar patterns (Supplementary Information, sections 2–5). Intriguingly, we also find that each of the three adsorbates desorbs at a different pressure, and that this adsorbate-specific desorption pressure is, to a first approximation, independent of the exact nature and pore size of the IRMOFs tested (Supplementary Tables 5, 9 and 11). This observation is another clear piece of evidence that adsorbate–adsorbate interactions within and across adjacent pores play a major role in gas uptake and release, both at the outset of the desorption process and in the formation of extra adsorption domains.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 22 June 2014; accepted 8 September 2015.
Published online 9 November 2015; corrected online 25 November 2015 (see full-text HTML version for details).

1. Kitagawa, S., Kitaura, R. & Noro, S. Functional porous coordination polymers. *Angew. Chem. Int. Edn* **43,** 2334–2375 (2004).
2. Furukawa, H., Cordova, K. E., O'Keeffe, M. & Yaghi, O. M. The chemistry and applications of metal-organic frameworks. *Science* **341,** 1230444 (2013).
3. Rosi, N. L. *et al.* Hydrogen storage in microporous metal-organic frameworks. *Science* **300,** 1127–1129 (2003).
4. Dincă, M. *et al.* Hydrogen storage in a microporous metal-organic framework with exposed Mn^{2+} coordination sites. *J. Am. Chem. Soc.* **128,** 16876–16883 (2006).
5. Farha, O. K. *et al. De novo* synthesis of a metal-organic framework material featuring ultrahigh surface area and gas storage capacities. *Nature Chem.* **2,** 944–948 (2010).
6. Holst, J. R. & Cooper, A. I. Ultrahigh surface area in porous solids. *Adv. Mater.* **22,** 5212–5216 (2010).
7. Makal, T. A., Li, J., Lu, W. & Zhou, H. Methane storage in advanced porous materials. *Chem. Soc. Rev.* **41,** 7761–7779 (2012).
8. Deng, H. *et al.* Large-pore apertures in a series of metal-organic frameworks. *Science* **336,** 1018–1023 (2012).
9. Nugent, P. *et al.* Porous materials with optimal adsorption thermodynamics and kinetics for CO_2 separation. *Nature* **495,** 80–84 (2013).
10. Rowsell, J. L. C., Spenser, E. C., Eckert, J., Howard, J. A. K. & Yaghi, O. M. Gas adsorption sites in a large-pore metal-organic framework. *Science* **309,** 1350–1354 (2005).
11. Vaidhyanathan, R. *et al.* Direct observation and quantification of CO_2 binding within an amine-functionalized nanoporous solid. *Science* **330,** 650–653 (2010).
12. Yang, S. *et al.* Selectivity and direct visualization of carbon dioxide and sulfur dioxide in a decorated porous host. *Nature Chem.* **4,** 887–894 (2012).
13. Serre, C. *et al.* Role of solvent-host interactions that lead to very large swelling of hybrid frameworks. *Science* **315,** 1828–1831 (2007).
14. Rabone, J. *et al.* An adaptable peptide-based porous material. *Science* **329,** 1053–1057 (2010).
15. Scherb, C., Koehn, R. & Bein, T. Sorption behavior of an oriented surface-grown MOF-film studied by in situ X-ray diffraction. *J. Mater. Chem.* **20,** 3046–3051 (2010).
16. Bureekaew, S. *et al.* Control of interpenetration for tuning structural flexibility impacts on sorption properties. *Angew. Chem. Int. Edn* **49,** 7660–7664 (2010).
17. Sato, H. *et al.* Self-accelerating CO sorption in a soft nanoporous crystal. *Science* **343,** 167–170 (2014).
18. Inagaki, S., Guan, S., Ohsuna, T. & Terasaki, O. An ordered mesoporous organosilica hybrid material with a crystal-like wall structure. *Nature* **416,** 304–307 (2002).
19. Zhao, D. *et al.* Triblock copolymer syntheses of mesoporous silica with periodic 50 to 300 angstrom pores. *Science* **279,** 548–552 (1998).
20. Joo, S. H. *et al.* Ordered nanoporous arrays of carbon supporting high dispersions of platinum nanoparticles. *Nature* **412,** 169–172 (2001).
21. Muroyama, N. *et al.* Argon adsorption on MCM-41 mesoporous crystal studied by in situ synchrotron powder X-ray diffraction. *J. Phys. Chem. C* **112,** 10803–10813 (2008).
22. Miyasaka, K., Neimark, A. V. & Terasaki, O. Density functional theory of in-situ synchrotron powder X-ray diffraction on mesoporouscrystals: argon adsorption on MCM-41. *J. Phys. Chem. C* **113,** 791–794 (2009).
23. Rosi, N. L. *et al.* Rod packings and metal-organic frameworks constructed from rod-shaped secondary building units. *J. Am. Chem. Soc.* **127,** 1504–1518 (2005).
24. Dietzel, P. D. C., Blom, R. & Fjellvåg, H. Base-induced formation of two magnesium metal-organic framework compounds with a bifunctional tetratopic ligand. *Eur. J. Inorg. Chem.* **2008,** 3624–3632 (2008).
25. Lin, L. C. *et al.* Understanding CO_2 dynamics in metal-organic frameworks with open metal sites. *Angew. Chem. Int. Edn* **52,** 4410–4413 (2013).
26. Bon, V. *et al.* In situ monitoring of structural changes during the adsorption on flexible porous coordination polymers by X-ray powder diffraction: instrumentation and experimental results. *Microporous Mesoporous Mater.* **188,** 190–195 (2014).
27. Gor, G. Y. *et al.* Adsorption of n-pentane on mesoporous silica and adsorbent deformation. *Langmuir* **29,** 8601–8608 (2013).

Supplementary Information is available in the online version of the paper.

Acknowledgements The authors acknowledge K. Ito, K. Sasaki, M. Kuribayashi and N. Muroyama (Rigaku America and Japan) and K. Nakai (Japan Bel) for technical support; N. Fujita and T. Nishimatsu (Tohoku University, Japan), H. Furukawa and Y. Zhang (University of California at Berkeley, USA) for their input; and A. Sawada (Kyoto University, Japan) for advice in designing the gas cell. Financial support was provided by WCU/BK21+ (to H.S.C., K.M., J.K.K., O.M.Y. and O.T.); HIMC of Global Frontier Project (2013M3A6B1078884) funded by the Ministry of Science, ICT and Future Planning and Korea Center for Artificial Photosynthesis (to J.K.K.); Berzelii Centre EXSELENT on Porous Materials (to O.T.); and BASF (Ludwigshafen, Germany) (to O.M.Y.). H.D. and Z.D. were supported by the 1000 Talent Plan of China, National Natural Science Foundation of China (21471118) and National Key Basic Research Program of China (2014CB239203). A.V.N. acknowledges support from the NSF ERC 'Structured Organic Particulate Systems'.

Author Contributions O.T., K.M. and J.K.K. designed and set up the experimental system. O.T. and O.M.Y. designed and led the project. H.S.C., K.M. and H.D. performed the SAXS experiments. H.D., Z.D. and M.C. prepared samples. A.V.N. contributed discussion of the gas adsorption–desorption process. H.D., H.S.C., J.K.K., O.M.Y. and O.T. prepared the first version of the manuscript and all authors contributed to the final version.

Additional Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to O.M.Y. (yaghi@berkeley.edu) or O. T. (terasaki@kaist.ac.kr).

METHODS

Synthesis of IRMOF-74 series. Organic linkers were synthesized as reported previously[8]. IRMOF-74 samples were synthesized by combining organic linkers with $Mg(NO_3)_2$ in a solution of dimethylformamide, ethanol and water, and then heated in an oven at 120 °C for 24 hours[8]. Needle-shaped crystals clustered in spherical forms were obtained. These IRMOF-74 samples were evacuated after solvent exchange with methanol nine times in three consecutive days to remove guest molecules.

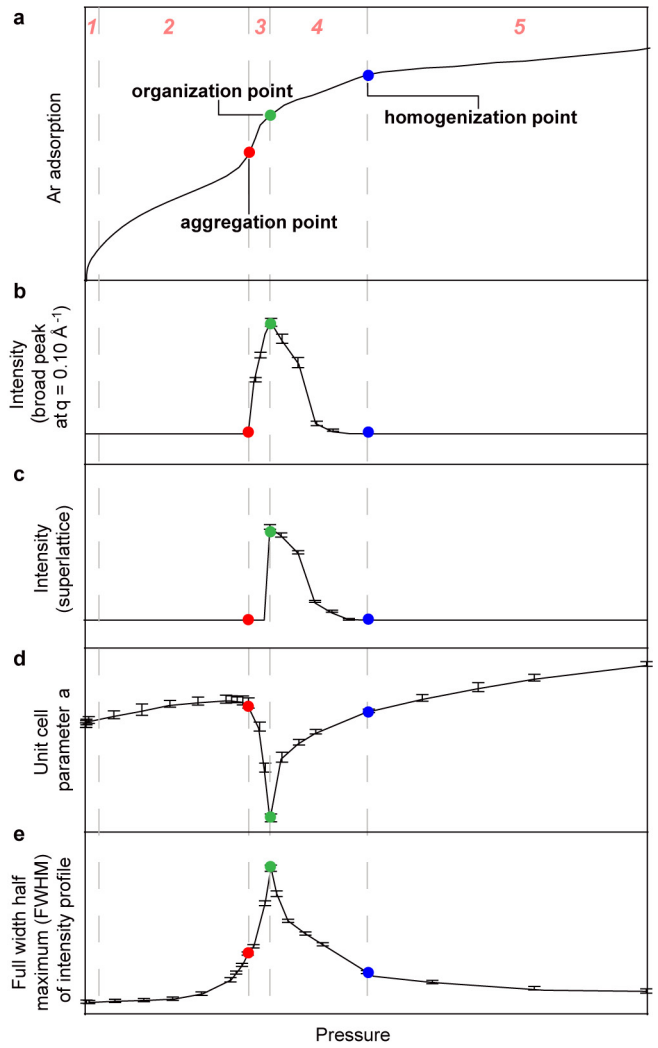
***In situ* gas adsorption SAXS measurement.** The *in situ* SAXS measurements for Ar, CO_2 and N_2 adsorption by IRMOF-74s (III, IV, V, VII and V-hex) were performed using a SAXS instrument (BioSAXS-1000; Rigaku, USA) equipped with a rotating anode X-ray source (FR-E+ Super Bright; Rigaku, Japan) and a gas adsorption instrument (BELSORP-max) together with a specially designed cell on a cryostat (Bel, Japan). We incorporated a sample cell inside the SAXS instrument, with a small chamber connected to the gas adsorption instrument placed outside. In addition, we used a large area detector combined with copper *K*α radiation from a rotating-anode X-ray source to provide precise measurement of both the intensity and the position of the diffraction peaks within a wide q ($=4\pi\sin\theta/\lambda$) range, from 0.01 to 0.71 $Å^{-1}$. Measurements were carried out with copper *K*α radiation in the transmission mode with Confocal Max Flux Mirror, a two-dimensional Kratky block and a Pilatus-type detector in the SAXS instrument. The powder samples were mounted in two places next to each other at the same adsorbate environmental condition: one was within the hollow part of the stainless steel rectangular plate covered by polyether ether ketone (PEEK) polymer films, in the X-ray path for diffraction; the other was for improving accuracy in measuring gas adsorption/desorption isotherms. The assembled samples were connected thermally to the temperature-controlling cryostat system, where the temperature is controlled within ±0.01 K, and to the gas adsorption instrument. The position of the sample cell was adjusted to the X-ray pathway within the chamber of the SAXS instrument at low temperatures before starting to take measurements.

A known weight (~0.03 g) of the IRMOF-74-III, IRMOF-74-IV, IRMOF-74-V, IRMOF-74-VII and IRMOF-74-V-hex samples was mounted in the sample cell and activated at 373 K for 6 hours under vacuum (~0.01 Pa) to remove the guest molecules before a series of measurements was taken. Activation of the IRMOF-74s was confirmed by comparing the argon isotherm of these MOFs at 87 K and CO_2 isotherm at 273 K with those of the activated sample measured on the traditional adsorption instrument (Supplementary Fig. 3). Gases (Ar, N_2 or CO_2) were introduced into the sample cell under measurement temperatures; the gas pressure was changed, and then maintained for 5 min after the system reached equilibrium (we judge the system to have reached equilibrium if the pressure fluctuation is less than 1 Pa for 5 min, which took roughly 30 min), for each measurement. The SAXS instrument was synchronized to the gas adsorption measurement and each SAXS pattern was collected at each equilibrium point of the sorption isotherms (the exposure time for each measurement was 30 minutes). There was no pressure change after the SAXS measurement, confirming that the sample with adsorbates in the sample cell was at equilibrium.

Before the actual adsorption/SAXS measurement started, gas adsorption without SAXS measurement was performed to confirm the adsorption curve and to set up the SAXS measurement points. We then collected SAXS scattering profiles at each of the 24 equilibrium points in the adsorption process, including the initial point (in vacuum). Another 21 profiles were collected for the desorption process. No transformation in the structure of the backbone of IRMOF-74 occurred throughout the whole gas adsorption process, as confirmed by the absence of obvious changes in peak positions in these SAXS patterns. Moreover, the samples did not show structural differences after *in situ* gas adsorption SAXS measurement, confirmed by adsorption data and SAXS data in the vacuum.

Structural analysis. For the structural analysis of IRMOF-74s at different gas pressures, Le Bail refinements[28] were performed using the JANA program[29] over the full sampled angular range, on the basis of the space group *R*3 for IRMOF-74-IV, IRMOF-74-V and IRMOF-74-VII, and $R\bar{3}$ for IRMOF-74-III and IRMOF-74-V-hex. The SAXS patterns of activated IRMOF-74 samples in the vacuum condition were refined first as a reference. The reflection peaks were modelled by a pseudo-Voigt peak–shape function modified for asymmetry, with six refinable coefficients. The background was treated using a Legendre polynomial with six refinable parameters. Because the q range for the SAXS instrument could cover only *hk*0 reflections owing to the small unit-cell parameter *c*, only unit-cell parameter *a* was refined. The standard deviations of all data were derived from comparison of observed points in SAXS profiles with corresponding ones calculated after Le Bail refinement. The atomic coordinates for all MOF samples were adopted from the framework structures derived from single-crystal X-ray-diffraction data[8] (Supplementary Tables 1–4). Because the number of reflections is limited, framework atomic coordinates were fixed for all data with different gas pressures. The distribution of adsorbates was calculated by difference Fourier analysis between the observed intensity and calculated intensity after a careful check of phase relationships among different reflections, and visualized using the VESTA program[30]. The calculated intensity was derived from the atomic coordinates obtained from single-crystal X-ray-diffraction analysis of MOF structures and the atomic coordinates were fixed for different gas pressures. The correct phase relationship of the crystal-structure factors between different reflections was verified by the fact that we could observe electrons at the open metal sites at the beginning of gas uptake. Electron-density-map data were illustrated using $\sqrt{3}\times\sqrt{3}$ *p*6 cell, which is the hexagonal projected structure of $R\bar{3}$ and *R*3, in order to show clearly the electron distribution in the pores (Fig. 1b, red parallelogram). The level of electron density (e^{-} $Å^{-3}$) is represented in blue/green/red colour code for all IRMOF-74 data. All electron-density-map data were presented with atomic coordinates of IRMOF-74 to clarify the relative position of adsorbates in the MOF.

28. Marra, G. L. *et al.* Cation location in dehydrated Na-Rb-Y zeolite: an XRD and IR study. *J. Phys. Chem. B* **101,** 10653–10660 (1997).
29. Petříček, V., Dušek, M. & Palatinus, L. Crystallographic computing system JANA2006: general features. *Zeitschrift Kristal. Crystalline Mater.* **229,** 345–352 (2006).
30. Momma, K. & Izumi, K. VESTA 3 for three-dimensional visualization of crystal, volumetric and morphology data. *J. Appl. Crystallogr.* **44,** 1272–1276 (2011).



Extended Data Figure 1 | The five stages of gas adsorption in IRMOF-74s. The five different adsorption stages are indicated in red at the top of the figure and their boundaries demarcated throughout all panels by grey dashed lines. **a**, The measured Ar adsorption by IRMOF-74-V-hex is shown; it can be compared against relevant SAXS profile features of IRMOF-74, measured as a function of Ar pressure, that are shown in the other panels. **b**, The appearance and disappearance of the broad peak indicates the formation of extra adsorption domains over pores (aggregation, red) and the even distribution of adsorbates (homogenization, blue). **c**, Intensity of $1\frac{1}{2}$ superlattice reflection, appearing as stage 3 turns to stage 4 (organization, green) and disappearing at the end of stage 4 (homogenization, blue). **d**, Change in the unit-cell parameter *a* of IRMOF-74. **e**, Change in the line-profile width of IRMOF-74.

LETTER

OPEN

doi:10.1038/nature15714

Single-molecule sequencing of the desiccation-tolerant grass *Oropetium thomaeum*

Robert VanBuren[1]*, Doug Bryant[1]*, Patrick P. Edger[2,3], Haibao Tang[4,5], Diane Burgess[2], Dinakar Challabathula[6]†, Kristi Spittle[7], Richard Hall[7], Jenny Gu[7], Eric Lyons[4], Michael Freeling[2], Dorothea Bartels[6], Boudewijn Ten Hallers[8], Alex Hastie[8], Todd P. Michael[9] & Todd C. Mockler[1]

Plant genomes, and eukaryotic genomes in general, are typically repetitive, polyploid and heterozygous, which complicates genome assembly[1]. The short read lengths of early Sanger and current next-generation sequencing platforms hinder assembly through complex repeat regions, and many draft and reference genomes are fragmented, lacking skewed GC and repetitive intergenic sequences, which are gaining importance due to projects like the Encyclopedia of DNA Elements (ENCODE)[2]. Here we report the whole-genome sequencing and assembly of the desiccation-tolerant grass *Oropetium thomaeum*. Using only single-molecule real-time sequencing, which generates long (>16 kilobases) reads with random errors, we assembled 99% (244 megabases) of the *Oropetium* genome into 625 contigs with an N50 length of 2.4 megabases. *Oropetium* is an example of a 'near-complete' draft genome which includes gapless coverage over gene space as well as intergenic sequences such as centromeres, telomeres, transposable elements and rRNA clusters that are typically unassembled in draft genomes. *Oropetium* has 28,466 protein-coding genes and 43% repeat sequences, yet with 30% more compact euchromatic regions it is the smallest known grass genome. The *Oropetium* genome demonstrates the utility of single-molecule real-time sequencing for assembling high-quality plant and other eukaryotic genomes, and serves as a valuable resource for the plant comparative genomics community.

The genomes of *Arabidopsis*[3], rice[4], poplar, grape and *Sorghum*[5] were first sequenced using high-quality and reiterative Sanger-based approaches producing a series of 'gold standard' reference genomes. The advent of next-generation sequencing (NGS) technologies reduced costs of sequencing substantially, which has enabled sequencing of over 100 plant genomes[1]. The quality of plant genome assemblies depends on genome size, ploidy, heterozygosity and sequence coverage, but most NGS-based genomes have on the order of tens of thousands of short contigs distributed in thousands of scaffolds. The short read lengths of NGS, inherent biases and non-random sequencing errors have resulted in highly fragmented draft genome assemblies that are not complete, which means they are missing biologically meaningful sequences including entire genes, regulatory regions, transposable elements, centromeres, telomeres and haplotype-specific structural variations. It is becoming clear from ENCODE projects that complete genomes are needed to better understand the importance of the non-coding regions of genomes[2].

More than 40% of calories consumed by humans are derived from grasses, and the grass family (Poaceae) is arguably the most important plant family with regard to global food security[6]. The size and complexity of most grass genomes has challenged progress in gene discovery and comparative genomics, although draft genomes are now available for most agriculturally important grasses[1]. The largest genome assemblies, such as maize (2,300 megabases (Mb))[7], barley (5,100 Mb)[8] and wheat (hexaploid, 17,000 Mb)[9] are highly fragmented as a result of the inability of current sequencing technologies to span complex repeat regions. Near-finished reference genomes are available for rice[4], *Sorghum*[5] and *Brachypodium*[10], but more high-quality grass genomes are needed for comparative genomics and gene discovery. Here we present the 'near-complete' draft genome of the grass *Oropetium thomaeum*, the first high-quality reference genome from the Chloridoideae subfamily. The draft genome is near complete because we were able to sequence through complex repeat regions that are unassembled in most draft genomes. *Oropetium* has the smallest known grass genome at 245 Mb and is also a resurrection plant that can survive the extreme water stress such as loss of >95% of cellular water (Fig. 1)[11].

Single-molecule real-time (SMRT) sequencing (Pacific Biosciences) produces long and unbiased sequences, which enables assembly of complex repeat structures and GC- and AT-rich regions that are often unassembled or highly fragmented in NGS-based draft genomes. We generated ~72× sequencing coverage of the *Oropetium* genome using 32 SMRT cells on the PacBio RS II platform (which is equivalent to <1 week of sequencing time and <US$10,000 in reagents). The resulting sequence had a read N50 length of over 16 kilobases (kb), and there was 10× coverage of reads over 20 kb in length (Extended Data Fig. 1a). The raw reads were error-corrected using the hierarchical genome assembly process (HGAP), and the longest reads (>16 kb) were assembled using Celera assembler followed by two rounds of genome polishing using Quiver[12]. The assembly contains 650 contigs spanning 99% (244 Mb) of the estimated 245 Mb genome size (Extended Data Fig. 1b) with a contig N50 length of 2.4 Mb (Extended Data Fig. 1c). The final assembly consists of 625 contigs after removal of the complete chloroplast genome, mitochondria-derived contigs and contaminants. The 35 largest contigs span half the genome, and the largest 107 contigs contain 90% of the sequence. The 135,324 base-pair (bp) chloroplast genome assembled into a single contig that includes both ~25 kb of inverted repeat regions which typically collapse into a single copy during assembly. The mitochondria genome was assembled into 20 partially overlapping circular chromosomes, which are the product of intramolecular recombination events that collectively span 1,100 kb.

The *Oropetium* genome has high contiguity for an uncurated draft plant genome. The average contig N50 length for all published plant genomes is 50 kb compared to 2.4 Mb for *Oropetium* (Extended Data Fig. 1d, e). After manual curation and data augmentation, only the *Arabidopsis* (TAIR10)[13], rice (V7) and *Brachypodium* (V 2.1)[10] genomes have longer contig N50 lengths. The accuracy rate is very

[1]Donald Danforth Plant Science Center, St Louis, Missouri 63132, USA. [2]Department of Plant and Microbial Biology, University of California Berkeley, Berkeley, California 94720, USA. [3]Department of Horticulture, Michigan State University, East Lansing, Michigan 48823, USA. [4]iPlant Collaborative, School of Plant Sciences, University of Arizona, Tucson, Arizona 85721, USA. [5]Center for Genomics and Biotechnology, Haixia Institute of Science and Technology (HIST), Fujian Agriculture and Forestry University, Fuzhou 350002, China. [6]IMBIO, University of Bonn, Kirschallee 1, D-53115 Bonn, Germany. [7]Pacific Biosciences, Menlo Park, California 94025, USA. [8]BioNano Genomics, San Diego, California 92121, USA. [9]Ibis Biosciences, Carlsbad, California 92008, USA.
†Present address: Department of Life Sciences, School of Basic and Applied Sciences, Central University of Tamil Nadu, Thiruvarur 610101, India.
*These authors contributed equally to this work.

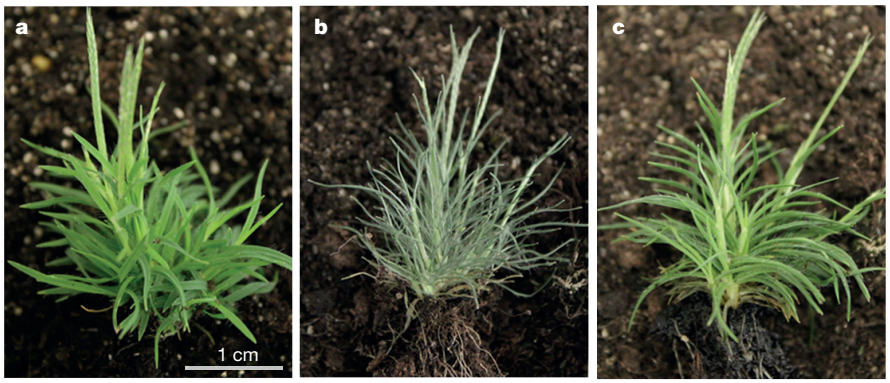


Figure 1 | Desiccation tolerance in the resurrection grass *Oropetium thomaeum*. **a**, Well watered. **b**, Desiccated (relative water content <5%) after 9 days of drought stress. **c**, Condition 24 h post-hydration (relative water content >70%).

high at 99.99995%, which is similar to Sanger-based approaches and higher than most NGS-based assemblies (Extended Data Fig. 1h). We plotted repeat density and GC content along the length of the contigs to identify factors causing contig breaks (Extended Data Fig. 1f, g). There is no correlation between repeat density and GC content at contig break points. This suggests that contig break points occur at the start of repeats or that most assembly breaks are caused by other factors, such as within-genome heterozygosity or haplotype-specific structural variation. To test this, we also tried 'diploid-aware' assemblers Falcon (https://github.com/PacificBiosciences/falcon) and MinHash Alignment Process (MHAP)[14]. These assemblies had similar metrics but were less contiguous overall (Extended Data Fig. 1i).

The completeness of the *Oropetium* genome allowed us to accurately survey its highly repetitive features that are often unassembled in most plant genomes. The *Oropetium* assembly captures all 18 telomeric arrays (Extended Data Table 1) with repeat number ranging from 40 to 900, suggesting that at least some are full length. Three of the nine centromeric satellites are completely assembled into large inverted repeats spanning 400 kb with a base monomer length of 155 bp, and higher order structures of dimers (310 bp), trimers (465 bp) and tetramers (620 bp; Fig. 2, Extended Data Fig. 2 and Supplementary Table 1). The remaining 40 centromeric sequences are incomplete centromere repeat fragments broken during assembly or solo repeats not associated with a larger centromere satellite. Nucleolus organizer regions contain tandem arrays of the 18S, 5.8S and 25S ribosomal RNA (rRNA) genes and typically span several megabase pairs with hundreds of nearly identical 10-kb arrays. Twenty-two full-length rRNA tandem arrays in six contigs are found in the *Oropetium* assembly (Extended Data Table 2). The largest tandem array contains five identical and one partial 9-kb repeats collectively spanning 51 kb; this is approaching the theoretical limit given the read-length distributions of our data. The remaining rRNA tandem repeats probably collapsed during read correction or genome assembly given their high sequence conservation.

Most repeats are incomplete, unassembled or highly collapsed in Illumina/454 NGS-based genomes, which has led to an underestimation and misclassification of repeat content in most plant genomes. Repetitive elements account for a surprisingly high proportion of the *Oropetium* genome (43%) compared to 21% in *Brachypodium*[10], 35% in rice[4], 54% in *Sorghum*[5] and over 90% in wheat[9] (Extended Data Table 3). Similar to these other genomes, the long terminal repeat (LTR) retrotransposons are the most abundant class and account for 35.6% of the *Oropetium* genome. We identified 3,247 intact LTRs in 358 families, which is similar to rice (3,663) and *Brachypodium* (2,162), but far less than *Sorghum* (17,022)[15]. Only ~2% of the repeats are unclassified, which reflects the completeness of individual repeat elements due to the long reads.

Genome size in the grasses varies by several orders of magnitude as a consequence of polyploidy and genome bloating due to repetitive DNA accumulation[16]. *Oropetium* has the smallest known genome among the grasses[17] at 90%, 60%, 50%, 30% and 10% the size of *Brachypodium*[10], rice[4], *Setaria*[18], *Sorghum*[5] and maize[7], respectively. We found that *Oropetium* has a solo:intact LTR ratio >1, which is similar to small grass genomes like rice and *Brachypodium*, where proliferating LTRs are removed by illegitimate recombination, whereas large grass genomes like *Sorghum* and maize have solo:intact LTR ratios <1 (ref. 15). Despite its compact size, the *Oropetium* genome has a typical number of predicted protein coding genes at 28,446. A pan-cereal whole-genome duplication (WGD) event, called rho, occurred before the diversification of grasses[5,19]. There appear to have been no further WGDs in the selected grass genomes, including *Oropetium*, since the shared rho event[4,5].

Genome alignments between *Oropetium* and selected grass genomes are mostly one-to-one after exclusion of the alignments derived from the shared genome duplication events (Extended Data Fig. 3a–e). Overall, 75% of the *Oropetium* genome, or 89% of its gene space, is contained in conserved syntenic blocks when compared to other grasses. Genomic colinearity across grass genomes is extensive, with a high density of orthologous genes spanning much of the euchromatin (Fig. 3). Insertions of retrotransposons and non-collinear genes that originated elsewhere in the genome contribute greatly to the differences in the intergenic sequences in grasses[20].

The relative sizes of syntenic blocks in the grass genomes track closely with the overall genome size difference (Extended Data Fig. 3f).

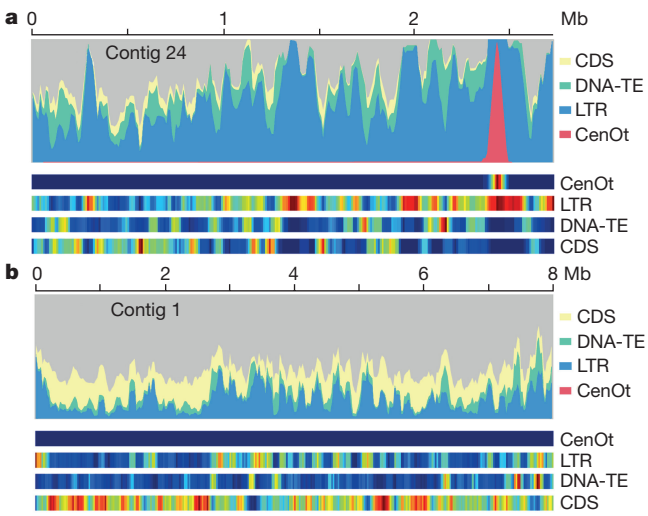


Figure 2 | SMRT sequencing enables contiguous sequencing over complex regions. The distributions of centromere-specific satellite DNA (CenOt), long terminal repeat retrotransposons (LTRs), DNA transposable elements (DNA-TE) and coding DNA sequences (CDS) are plotted. **a**, The gap-free assembly of a full-length centromeric array and the flanking highly repetitive pericentromeric region. **b**, The largest contig (7.8 Mb), which has a more typical distribution of elements.

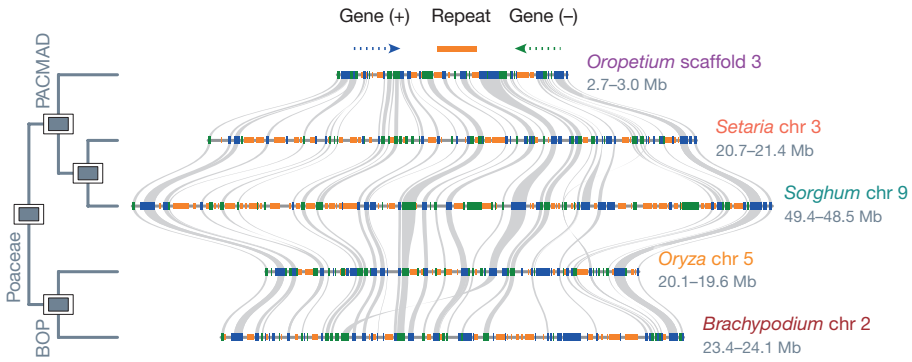


Figure 3 | Compact genome structure of *Oropetium*. *Oropetium*, part of the PACMAD clade, provides the first high-quality reference genome from the Chloridoideae subfamily—a large and diverse group of ~1,600 species that contains the orphan crops tef (*Eragrostis tef*) and finger millet (*Eleusine coracana*). Typical micro-colinearity patterns among genomic regions from *Oropetium*, *Setaria*, *Sorghum*, *Oryza* and *Brachypodium* are shown. Rectangles show predicted gene models, and colours indicate relative orientations. Matching gene pairs are displayed as grey connections. chr, chromosome.

In contrast, the genomic span of coding sequences is similar across genes that are retained in orthologous locations, although coding features are slightly smaller in *Oropetium* (Extended Data Table 4). The relatively constant sizes of coding sequences among grass genomes confirm that genome size differences are indeed due to variations in the intergenic contents. It was thought that plants have a 'one-way ticket to genome obesity' due to the retention of proliferating transposable elements[21]. However, analysis of carnivorous plants *Utricularia gibba* (bladderwort, 82 Mb)[22] and *Genlisea aurea* (corkscrew, 63.6 Mb)[23] provided evidence that almost all intergenic space can be purged. Small genomes also arise from a reduction in gene number as seen in the aquatic monocotyledon *Spirodela polyrhiza*, which has the fewest predicted protein coding genes at 19,623 (ref. 24). *Oropetium* seems to have reduced both its intergenic and intragenic sequence.

As the intergenic sequence in *Oropetium* is specifically reduced compared with other grasses (Extended Data Fig. 3f), we determined which sequence accounted for its smaller genome size by comparing highly syntenic regions of the larger 730 Mb *Sorghum* genome. To identify highly orthologous regions we looked for *Sorghum* genes (promoter, 5′UTR, exons, introns and 3′UTR) with an increased number of conserved noncoding sequences[25]. We then analysed the top 48 *Sorghum* genes against their orthologous sequences in *Oropetium* and found that they were 38% (±0.27, 1 s.d.) larger in *Sorghum* (Extended Data Fig. 4a). The primary driver of gene-space expansion was highly unique ~1-kb intragenic sequences evenly spaced within the *Sorghum* genes. One explanation is that these evenly spaced highly unique sequences are degenerate remnants of transposons that have been partly purged from the *Sorghum* genome. *Oropetium* has a >1 solo:intact LTR ratio, consistent with active purging of transposons and complete loss of these regions. These results lend support to an emerging theory about the C-value paradox called the Genome Balance Hypothesis[26], which suggests that selection on gene networks and pericentromeric growth (centromere movement) is balanced by transposon proliferation and retention. Therefore, these evenly spaced highly unique sequences balance the 6:1 expansion of pericentromeric sequence in *Sorghum* as compared to *Oropetium* (Extended Data Fig. 4b).

Desiccation tolerance was a key adaptation that permitted the most recent common ancestor of terrestrial plants to survive on land. Desiccation tolerance is widespread in bryophytes and lichens but rare in flowering plants, although similar mechanisms have evolved in vascular plants for seed and pollen desiccation. Desiccation tolerance to survive prolonged drought evolved independently in diverse monocotyledon and eudicotyledon lineages, and is found in at least 300 species. Gene duplications have provided the raw material for evolutionary innovation across plants. Tandem duplicated genes are often involved in stress responses and are probably important for adaptive evolution in dynamically changing environments. *Oropetium* has 6,668 tandem duplicated genes in 2,326 clusters, which is a slightly higher number than in other grasses, but a similar proportion (24% of genes). Tandem duplicated genes are enriched for gene ontology terms involved in response to abiotic stresses, gene regulation and cellular metabolism (Supplementary Table 2). In addition, *Oropetium* has 4,209 homeologous gene pairs retained from the rho WGD event, which are enriched for gene ontology terms related to gene regulation and stress responses such as transcription factor activity, nitrogen metabolism, response to abiotic stimulus, to salt stress and to oxygen-containing compounds (Supplementary Tables 3 and 4). Understanding the genomic mechanisms of extreme desiccation tolerance in resurrection plants such as *Oropetium* may provide targets for engineering drought and stress tolerance in crop plants.

Pacific Biosciences (PacBio) SMRT sequencing has been used to close gaps in the human genome[27], assemble complete bacterial genomes[12] and identify novel gene isoforms[28]. Here we present a several hundred megabase plant genome, sequenced and assembled entirely by SMRT sequencing. The long SMRT reads produced a near-complete draft genome that captured three of nine complete centromeres, all of the telomeres and biologically relevant features of the *Oropetium* genome. The total time from extracted DNA to a complete assembly was less than one month, and costs for PacBio were comparable to an Illumina-based genome assembly. Our study demonstrates that SMRT sequencing enables a new level of genome assembly required for full ENCODE-type analysis of intergenic sequence, which is not currently possible with other NGS-based methods. The compactness of the *Oropetium* genome results from purging of both inter- and intragenic sequences, probably through small deletions during illegitimate recombination, as has been shown in other grasses. One hypothesis is that genome size is a function of cell size[29], and consistent with this, all small plant genomes sequenced to date including *Arabidopsis* (125 Mb), *Brachypodium* (272 Mb), *Selaginella* (100 Mb) *Spirodela* (158 Mb) and *Utricularia* (82 Mb) are plants of very small stature (Fig. 1). However, we provide evidence for the Genome Balance Hypothesis, which suggests that there is selective pressure on *Oropetium* to purge proliferating transposons in order to maintain expression balance of networked genes and spacing in centromeres. The complete assembly of complex and highly similar repeat sequences demonstrated here suggests that SMRT sequencing can be used to assemble large and polyploid plant and other eukaryotic genomes, assuming ample sequence coverage and computational resources. SMRT-sequencing-based assemblies provide an opportunity to determine how these regions play a role in genome architecture and dynamics.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 28 April; accepted 10 September 2015.
Published online 11 November 2015; corrected online 25 November 2015 (see full-text HTML version for details).

1. Michael, T. P. & VanBuren, R. Progress, challenges and the future of crop genomes. *Curr. Opin. Plant Biol.* **24,** 71–81 (2015).
2. Kellis, M. *et al.* Defining functional DNA elements in the human genome. *Proc. Natl Acad. Sci. USA* **111,** 6131–6138 (2014).
3. The Arabidopsis Genome Initiative. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408,** 796–815 (2000).
4. International Rice Genome Sequencing Project. The map-based sequence of the rice genome. *Nature* **436,** 793–800 (2005).
5. Paterson, A. H. *et al.* The *Sorghum bicolor* genome and the diversification of grasses. *Nature* **457,** 551–556 (2009).
6. Elert, E. Rice by the numbers: A good grain. *Nature* **514,** S50–S51 (2014).
7. Schnable, P. S. *et al.* The B73 maize genome: complexity, diversity, and dynamics. *Science* **326,** 1112–1115 (2009).
8. International Barley Genome Sequencing Consortium. A physical, genetic and functional sequence assembly of the barley genome. *Nature* **491,** 711–716 (2012).
9. International Wheat Genome Sequencing Consortium (IWGSC). A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science* **345,** 1251788 (2014).
10. The International Brachypodium Initiative. Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature* **463,** 763–768 (2010).
11. Bartels, D. & Mattar, M. *Oropetium thomaeum*: A resurrection grass with a diploid genome. *Maydica* **47,** 185–192 (2002).
12. Chin, C.-S. *et al.* Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nature Methods* **10,** 563–569 (2013).
13. Lamesch, P. *et al.* The *Arabidopsis* Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Res.* **40,** D1202–D1210 (2012).
14. Berlin, K. *et al.* Assembling large genomes with single-molecule sequencing and locality-sensitive hashing. *Nature Biotechnol.* **33,** 623–630 (2015).
15. El Baidouri, M. & Panaud, O. Comparative genomic paleontology across plant kingdom reveals the dynamics of TE-driven genome evolution. *Genome Biol. Evol.* **5,** 954–965 (2013).
16. Michael, T. P. Plant genome size variation: bloating and purging DNA. *Brief. Funct. Genomic.* **13,** 308–317 (2014).
17. Jones, N. & Pašakinskienė, I. Genome conflict in the gramineae. *New Phytol.* **165,** 391–410 (2005).
18. Bennetzen, J. L. *et al.* Reference genome sequence of the model plant Setaria. *Nature Biotechnol.* **30,** 555–561 (2012).
19. Tang, H., Bowers, J. E., Wang, X. & Paterson, A. H. Angiosperm genome comparisons reveal early polyploidy in the monocot lineage. *Proc. Natl Acad. Sci. USA* **107,** 472–477 (2010).
20. Wicker, T., Buchmann, J. P. & Keller, B. Patching gaps in plant genomes results in gene movement and erosion of colinearity. *Genome Res.* **20,** 1229–1237 (2010).
21. Bennetzen, J. L. & Kellogg, E. A. Do plants have a one-way ticket to genomic obesity? *Plant Cell* **9,** 1509 (1997).
22. Ibarra-Laclette, E. *et al.* Architecture and evolution of a minute plant genome. *Nature* **498,** 94–98 (2013).
23. Leushkin, E. V. *et al.* The miniature genome of a carnivorous plant *Genlisea aurea* contains a low number of genes and short non-coding sequences. *BMC Genomics* **14,** 476 (2013).
24. Wang, W. *et al.* The *Spirodela polyrhiza* genome reveals insights into its neotenous reduction fast growth and aquatic lifestyle. *Nature Commun.* **5,** 3311 (2014).
25. Lyons, E. & Freeling, M. How to usefully compare homologous plant genes and chromosomes as DNA sequences. *Plant J.* **53,** 661–673 (2008).
26. Freeling, M., Xu, J., Woodhouse, M. & Lisch, D. A solution to the C-value paradox and the function of junk DNA: the Genome Balance Hypothesis. *Mol. Plant* **8,** 899–910 (2015).
27. Chaisson, M. J. P. *et al.* Resolving the complexity of the human genome using single-molecule sequencing. *Nature* **517,** 608–611 (2015).
28. Au, K. F. *et al.* Characterization of the human ESC transcriptome by hybrid sequencing. *Proc. Natl Acad. Sci. USA* **110,** E4821–E4830 (2013).
29. Beaulieu, J. M., Leitch, I. J., Patel, S., Pendharkar, A. & Knight, C. A. Genome size is a strong predictor of cell size and stomatal density in angiosperms. *New Phytol.* **179,** 975–986 (2008).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work is supported in part by funding from the National Science foundation (DBI-1401572 to R.V.; DBI-120793 to P.P.E.), USDA NIFA (CO471A-B to M.F.), the Department of Energy (DE-SC0012639 to T.C.M. and T.P.M.; DE-SC-0008769 to T.C.M.), the Donald Danforth Plant Science Center to T.C.M. and the Enterprise Rent-A-Car Institute for Renewable Fuels to T.C.M. Sequencing was provided by Pacific Biosciences under the 'Most Interesting Genome in the World' 2014 SMRT grant program.

Author Contributions R.V., D.Br., T.P.M. and T.C.M. designed and conceived research; D.Ba. and D.C. identified biological material, performed desiccation experiments and extracted DNA and RNA; R.V. prepared DNA for PacBio sequencing; T.P.M., R.V. and T.C.M. performed Illumina sequencing; K.S., R.H. and J.G. performed PacBio sequencing and assembly; B.T.H. and A.H. conducted the BioNano analysis. D.Br., R.V., T.P.M. and T.C.M. annotated genome features; E.L., M.F., D.Bu., R.V., D.Br., H.T., T.P.M., T.C.M. and P.P.E. analysed data; R.V., T.P.M. and T.C.M. wrote the paper. All authors read and approved the final manuscript.

Author Information The genome assembly and annotation have been deposited in CoGe under the accession code 25799 (https://genomevolution.org/CoGe/GenomeInfo.pl?gid=25799), in the NCBI BioProject under PRJNA286116, and in GenBank under accession number LFJQ00000000. Raw PacBio and Illumina reads are available at the Short Read Archive at NCBI under the aforementioned NCBI BioProject. Genome assembly and annotation are also available at http://www.oropetium.org/. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.P.M. (toddpmichael@gmail.com) or T.C.M. (tmockler@danforthcenter.org).

METHODS

No statistical methods were used to predetermine sample size.

Plant material. *Oropetium thomaeum* is a compact resurrection plant that has the smallest known genome among the grasses, at 245 Mb and 9 chromosomes ($2n = 2x = 18$; 1C = 0.25 pg)[17]. We estimated the genome size to be 250 Mb by flow cytometry and 245 Mb by *k*-mer analysis (Extended Data Fig. 1b). *Oropetium thomaeum* plants were originally collected in Jodhpur, Rajasthan, India and propagated as previously described[11]. *Oropetium* is a member of the Chloridoideae subfamily, a large and diverse group of roughly 1,600 species that contains the orphan crops tef (*Eragrostis tef*) and finger millet (*Eleusine coracana*) as well as some turf grasses (such as Bermuda grass, *Cynodon dactylon* and *Zoysia japonica*).

SMRT PacBio sequencing. Fifty micrograms of high-molecular-weight *Oropetium* gDNA was extracted using a modified nuclei preparation method[30] followed by an additional high-salt phenol–chloroform purification to minimize contamination. A 20-kb insert SMRTbell library was generated using a 15 kb lower-end size selection protocol on the BluePippin (Sage Science). Initial titration runs were performed to optimize loading on the SMRT Cell for maximum performance. The *Oropetium* genome was sequenced using 32 SMRT Cells with 4-h collections and P6-C4 chemistry on the PacBio RS II platform (Pacific Biosciences).

HGAP genome assembly. The *Oropetium* genome was assembled using the RS_HGAP_Assembly.3 protocol for assembly and Quiver for genome polishing in SMRT Analysis v2.3.0[12]. This consisted of a three-step process involving (1) generation of preassembled reads with improved consensus accuracy; (2) assembly of the genome through overlap consensus accuracy using Celera; and (3) one round of genome polishing with Quiver. For HGAP, the following parameters were used: PreAssembler Filter v1 (minimum sub-read length = 3,000 bp, minimum polymerase read quality = 0.80, minimum polymerase read length = 3,000 bp); PreAssembler v2 (minimum seed length = 16,000 bp, number of seed read chunks = 6, alignment candidates per chunk = 10, total alignment candidates = 24, min coverage for correction = 6); AssembleUnitig v1 (target genome coverage = 30, overlap error rate = 0.06, minimum overlap = 40 bp and overlap *k*-mer = 14); and BLASR v1 mapping of reads for genome polishing with Quiver (max divergence percentage = 30, minimum anchor size = 12). A second round of genome polishing was performed using Quiver (SMRT Analysis v2.3.0) to further improve the site-specific consensus accuracy of the assembly. The following Quiver parameters were used for genome polishing: filtering (minimum sub-read length = 3,000 bp, minimum polymerase read quality = 0.80, minimum polymerase read length = 3,000 bp); mapping (maximum divergence percentage = 30, minimum anchor size = 12). Default parameters were otherwise employed for both HGAP assembly and Quiver protocols.

Falcon and MHAP assemblies. We also tested other assemblers to compare the PacBio HGAP assembly results (Extended Data Fig. 1i). Raw PacBio reads were error-corrected and assembled using Falcon and MHAP under default parameters. The Falcon and MHAP assemblies have lower contiguity than the HGAP assembly and have fewer assembled centromere and telomere sequences with a lower average length.

Construction of a genome map using the Irys system for contig anchoring and scaffolding. Genome mapping from BioNano Genomics[31] was used to improve the assembly quality of the *Oropetium* genome with the eventual goal of producing a chromosome-scale assembly. High molecular weight genomic DNA was isolated from fresh *Oropetium* tissue using the following protocol outline. Three grams of leaves were collected from live *Oropetium thomaeum* plants and fixed with formaldehyde. After blending with a tissue homogenizer in isolation buffer, a filtration step and Triton-X washing treatment were performed. The nuclei were purified on percoll cushions. The nuclei were washed extensively and embedded in low melting agarose at different dilutions. Finally, the DNA plugs were treated with a lysis buffer containing detergent, proteinase K and β-mercaptoethanol (BME). In total, 53 Gb of data (>100 kb) were collected representing ~200× genome coverage with a molecule N50 length of 169 kb (Extended Data Fig. 5a). The size distribution was lower than expected and is probably a result of impurities during high-molecular-weight gDNA isolation that would cause shearing and inhibition of enzymes. Molecules were *de novo* assembled as previously described[32]. Two genome maps were assembled at different stringencies, map set 1 has 402 maps with an N50 length of 725 kb and spans 216 Mb (Extended Data Fig. 5b); the second genome map has 214 maps and an N50 of 1.674 Mb. Combining the genome maps with the PacBio assembly to produce a hybrid scaffold was performed sequentially with the two genome maps. The scaffolding merged 90 contigs producing an assembly of 46 primary scaffolds covering 94% of the sequence assembly with an N50 of 7.8 Mb; in total there are 535 scaffolds with an N50 of 7.1 Mb and total assembled size of 244 Mb.

Variant calling using Illumina data. WGS Illumina sequences from *Oropetium* gDNA were used to assess the error rate of the PacBio assembly and residual within-genome heterozygosity (Supplementary Table 5). Raw Illumina HiSeq data from three different libraries of 570-bp insert, 1-kb insert and 3-kb insert sizes were trimmed for quality using Trimmomatic (v.0.32; ref. 33). Illumina sequence adaptors were removed, leading low quality (below quality 3) and N base pairs were trimmed, and reads were scanned using a 4-bp sliding window and trimmed when the average quality per base dropped below 30. Read pairs where both reads were ultimately of at least 36 bp in length following this quality control process were retained and used for subsequent analyses.

Quality trimmed data were aligned to our assembly using BWA mem (v. 0.7.12-r1039)[34]. Duplicate alignments were marked using Picard tools v.1.104 MarkDuplicates (http://broadinstitute.github.io/picard/). Genome Analysis Toolkit (v.3.3.0)[35] IndelRealigner was used to perform local realignment around indels, followed by application of GATK HaplotypeCaller to call variants. Identified single nucleotide polymorphisms were filtered by depth, strand bias, mapping quality and read position. Identified indels were filtered by depth, strand bias and read position.

The native error rate of raw PacBio reads is in the range of 15–20%, raising the possibility that residual sequencing errors may be introduced into the final assembly of the *Oropetium* genome. Homozygous mismatches are classified as sequencing errors, and heterozygous mismatches indicate sites of heterozygosity. The accuracy rate is very high at 99.99995%, and a relatively high proportion of the errors (two-thirds) are small insertions or deletions (indels). The accuracy rate is similar to those obtained with WGS Sanger approaches[5,36] and is higher than those reported for most NGS-based assemblies. The estimated residual within-genome heterozygosity for the *Oropetium* genome is very low at 0.087%, which probably contributed to the high contiguity of the assembly. This suggests that provided sufficient coverage, a PacBio SMRT-only approach can produce a high-quality complete plant genome.

Repeat annotation. To structurally annotate repeat sequences in the *Oropetium* genome, we began by discovering repetitive elements through application of the REPET v.2.2 packages TEdenovo and TEannot[37]. The TEdenovo pipeline compares the genome with itself to identify and classify repeated genomic elements. All-by-all alignments were conducted with NCBI-BLAST+ using default TEdenovo parameters. LTRharvest[38] was used for structural detection. During clustering, Grouper, Recon and Plier steps were invoked both with and without structural detection. Consensus building was performed using default parameters. During consensus detect features, repeat scout[39] was invoked, and Pfam26.0 HMM profiles[40] and Repbase (v18.08) nucleotide and amino acid databanks were used. Finally, consensus classification, filtering and clustering were performed using default parameters.

Output from the TEdenovo pipeline was used as input to the TEannot pipeline. This pipeline mines the genome sequence using repeated sequences identified in the previous TEdenovo pipeline to produce classified non-redundant consensus repeat sequences along with short simple repeats, which are exported to GFF3 format. First, a set of perfectly matching sequences from the TEdenovo-output transposable elements (TE) library was selected by running a subset of the TEannot pipeline, producing a working reference TE library. This TE library was used in a full run of the TEannot pipeline. For alignment of the reference TE library, NCBI-BLAST+ was used, and blaster, repeat masker and censor steps were run both on the reference TE library and on randomized chunks. Filtering was applied using default parameters. Short simple repeats were identified using the crossmatch engine. Merging was performed using default parameters. For comparisons, Repbase (v18.08) nucleotide and amino acids databanks were used. Finally, filtering was applied using default parameters, and annotations were exported to GFF3 format.

To classify identified repeats, non-redundant consensus repeat sequences as output by TEanno were annotated via PASTEClassifier v1.0 https://urgi.versailles.inra.fr/Tools/PASTEClassifier/README). To classify these sequences, Repbase (v18.08)[41] nucleotide and amino acid sequences were used, as were Pfam v26.0 (http://pfam.xfam.org/) HMM repeat profiles. Finally, identified LTRs were classified as Gypsy if homology or motif evidence existed for Gypsy and not for Copia, classified as Copia if the opposite were true, and otherwise classified as unknown.

Centromere and telomere identification. Centromeric repeats were identified using an approach outlined in ref. 42. Tandem repeat finder (TRF, Version 4.07b)[43] was used to find tandem repeats using the parameters '1 1 2 80 5 200 2000-d –h' in order to find high order repeats. The resulting '.dat file' was transformed into a GFF3 file, which was used to identify telomeric and centromeric repeats. To identify the centromeric repeats, the largest repeat arrays (period length X copy number) were identified and clustered. Clustered centromeric repeat regions were transformed into FASTA files and aligned using clustalX to identify array sequence composition and orientation. The base centromere repeat was 155 bp dimers (310 bp), trimers (465 bp) and tetramers (620 bp) (Extended Data Fig. 2 and Supplementary Table 1). The three largest centromeric arrays (contigs 003, 028 and 064) were >400 kb and resolved into large inverted repeats, consistent with them being full length. The telomeric repeats were identified by searching

the ends of contigs for short (~7 bp) high copy number repeats; 18 telomeric repeat sequences with the monomer 'AAACCCT' were identified (Extended Data Table 1).

Transcriptome assembly. Total RNA was extracted from fresh, desiccated and 24-h post rehydration *Oropetium* leaf tissues with 2 biological replicates collected for each tissue. RNA-seq libraries were prepared from the total RNA and bar-coded using TruSeq RNA Sample Prep Kits (Illumina) according to the manufacturer's protocol. Raw Illumina RNA-seq data from the six libraries were trimmed for quality using Trimmomatic (v.0.32; ref. 33). Illumina sequence adaptors were removed, then leading low-quality (below quality 3) and N base pairs were trimmed and, finally, resulting trimmed reads were scanned using a 4-bp sliding window and cut when the average quality per base dropped below 30. Read pairs where both reads were ultimately of at least 36 base pairs in length following this quality control process were retained and used for subsequent analyses. Trinity (v.r20140717)[44] was used to assemble quality filtered data. Assembled transcripts were aligned to our genome sequence using NCBI blastn v.2.2.30+ with an *e*-value cut-off of 1×10^{-5}. Successfully aligned transcripts were clustered at 90% identity using CD-HIT (v. 4.5.4)[45], with representative sequences from each cluster retained and used to help parameterize gene calling. Eighty-seven per cent of the trimmed RNA-seq reads aligned to the *Oropetium* genome, suggesting that the genome is largely complete (Supplementary Table 5). Reads that failed to align may have been contaminants from other organisms.

Gene annotation. Maker v2.31.8[46] (http://www.yandell-lab.org/software/maker.html) was used to identify putative genes. Aligned and representative sequences from our transcriptome assembly were input to Maker as expressed sequence tag evidence. Rice and *Brachypodium* proteome sequences clustered at 90% identity using CD-HIT (v. 4.5.4)[45] with representative sequences from each cluster retained and input to Maker as multi-organismal protein homology evidence. The *Oropetium* repeat database was input to Maker as a custom repeat library. SNAPhmm, Augustus, and GeneMarkHMM were invoked by Maker and were initially trained using rice and maize. Only genes for which the encoded protein was predicted to contain a complete open reading frame were retained.

On the basis of the gene annotations provided by Maker, cufflinks (v2.2.1)[47] was used to identify predicted genes without empirical expression evidence. Quality-trimmed data from all six RNA-seq libraries were input simultaneously to cufflinks, with results used to identify genes with and without expression.

Protein sequences from genes predicted by Maker were functionally annotated using NCBI blastp v.2.2.30+ versus the NCBI non-redundant refseq protein database (http://www.ncbi.nlm.nih.gov/refseq/), versus the UniProt database[48], and using InterProScan (v. 5.6-48.0)[49].

Finally, Maker-predicted genes were pruned based on a Maker-defined annotation edit distanced (AED) score that measures distance between the predicted gene and the evidence input to Maker, non-redundant (NR) annotation, Uniprot annotation, InterProScan annotation and expression level as output by cufflinks. Genes were removed that had no alignment evidence (AED = 1), no sequence match to either the NR or Uniprot databases, no InterProScan predicted domains and no expression evidence in our RNA-seq data.

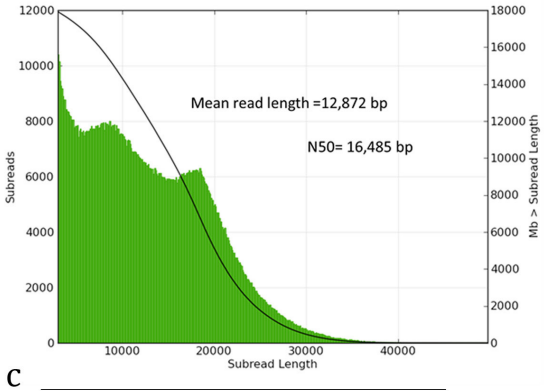
Synteny and comparative genomics. Genome data sets from *Setaria*, *Sorghum*, rice and *Brachypodium* were downloaded from Phytozome (version 9.1) and subject to pairwise genome alignments against the *Oropetium* genome. For each pairwise alignment, the coding sequences of predicted gene models are compared to each other using adaptive seeds[50]. Our synteny search pipeline defines syntenic blocks by chaining the large-scale alignment tool (LAST) hits with a distance cut-off of 20 genes apart, also requiring at least four gene pairs per syntenic block. The syntenic blocks were further screened using QUOTA-ALIGN[51] to retain one-to-one blocks and to exclude weak blocks derived from shared ancient duplications. The resulting dot plots were visually inspected to confirm the structural similarity of the *Oropetium* genome in relation to other genomes (Extended Data Fig. 3a–e).

Pairwise genomic alignments, described above, combined with OrthoMCL[52] analyses filtered to one-to-one hits were used to identify orthologous gene clusters between *Oropetium* and *Sorghum*, rice, *Vitis* and *Arabidopsis*. The complete *Oropetium–Arabidopsis* orthologue list was then filtered to focus on genes with functional data in the STRING v9.1 global *Arabidopsis* protein interaction network[53]. Gene expression patterns and duplicated genes (tandem and whole-genome duplicates) were mapped onto this network using Cytoscape v3.1.1[54] to identify clusters of co-expressed and interacting duplicate genes, respectively (Extended Data Fig. 6). Various network statistics were calculated using NetworkAnalyzer[55], including average number of neighbours (that is, protein interactions) and total number of isolated nodes (that is, without known interactors).

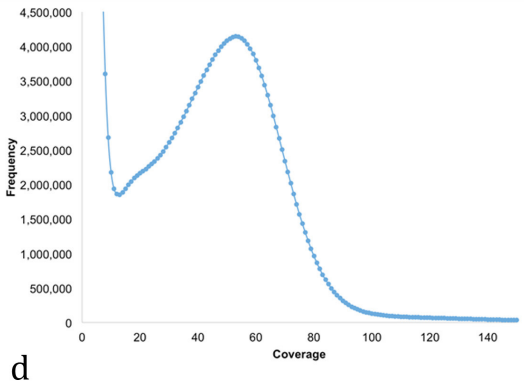
Constructing a gene interaction network. We constructed a gene interaction network for *Oropetium* on the basis of orthologous relationships with *Arabidopsis* genes with validated interactions and expression data yielding a network with 4,421 nodes (gene products) with 36,918 edges (interactions). This network encompasses most metabolic pathways including photosynthesis, core anabolic and catabolic processes and stress response pathways (Extended Data Fig. 6).

30. Zhang, H.-B., Zhao, X., Ding, X., Paterson, A. H. & Wing, R. A. Preparation of megabase-size DNA from plant nuclei. *Plant J.* **7,** 175–184 (1995).
31. Lam, E. T. *et al.* Genome mapping on nanochannel arrays for structural variation analysis and sequence assembly. *Nature Biotechnol.* **30,** 771–776 (2012).
32. Cao, H. *et al.* Rapid detection of structural variation in a human genome using nanochannel-based genome mapping technology. *GigaScience* **3,** 34 (2014).
33. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30,** 2114–2120 (2014).
34. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25,** 1754–1760 (2009).
35. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20,** 1297–1303 (2010).
36. Ming, R. *et al.* The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya Linnaeus*). *Nature* **452,** 991–996 (2008).
37. Flutre, T., Duprat, E., Feuillet, C. & Quesneville, H. Considering transposable element diversification in de novo annotation approaches. *PLoS ONE* **6,** e16526 (2011).
38. Ellinghaus, D., Kurtz, S. & Willhoeft, U. LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. *BMC Bioinformatics* **9,** 18 (2008).
39. Price, A. L., Jones, N. C. & Pevzner, P. A. De novo identification of repeat families in large genomes. *Bioinformatics* **21,** i351–i358 (2005).
40. Finn, R. D. *et al.* Pfam: the protein families database. *Nucleic Acids Res.* **42,** D222–D230 (2014).
41. Jurka, J. *et al.* Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet. Genome Res.* **110,** 462–467 (2005).
42. Melters, D. P. *et al.* Comparative analysis of tandem repeats from hundreds of species reveals unique insights into centromere evolution. *Genome Biol.* **14,** R10 (2013).
43. Benson, G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* **27,** 573 (1999).
44. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnol.* **29,** 644–652 (2011).
45. Huang, Y., Niu, B., Gao, Y., Fu, L. & Li, W. CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics* **26,** 680–682 (2010).
46. Cantarel, B. L. *et al.* MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* **18,** 188–196 (2008).
47. Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* **7,** 562–578 (2012).
48. Wu, C. H. *et al.* The Universal Protein Resource (UniProt): an expanding universe of protein information. *Nucleic Acids Res.* **34,** D187–D191 (2006).
49. Quevillon, E. *et al.* InterProScan: protein domains identifier. *Nucleic Acids Res.* **33,** W116–W120 (2005).
50. Kiełbasa, S. M., Wan, R., Sato, K., Horton, P. & Frith, M. C. Adaptive seeds tame genomic sequence comparison. *Genome Res.* **21,** 487–493 (2011).
51. Tang, H. *et al.* Screening synteny blocks in pairwise genome comparisons through integer programming. *BMC Bioinformatics* **12,** 102 (2011).
52. Li, L., Stoeckert, C. J. & Roos, D. S. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* **13,** 2178–2189 (2003).
53. Franceschini, A. *et al.* STRING v9. 1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res.* **41,** D808–D815 (2013).
54. Saito, R. *et al.* A travel guide to Cytoscape plugins. *Nature Methods* **9,** 1069–1076 (2012).
55. Doncheva, N. T., Assenov, Y., Domingues, F. S. & Albrecht, M. Topological analysis and interactive visualization of biological networks and protein structures. *Nature Protocols* **7,** 670–685 (2012).

a



b



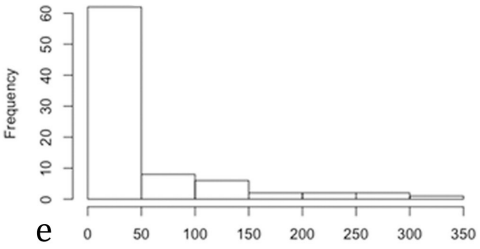
c

Raw Input:	
Mean Subread Length	12,872 bp
N50 (Subread Length)	16,485 bp
Total Number of sequenced Bases	18,022,966,707 bp
Number of Reads	1,400,150

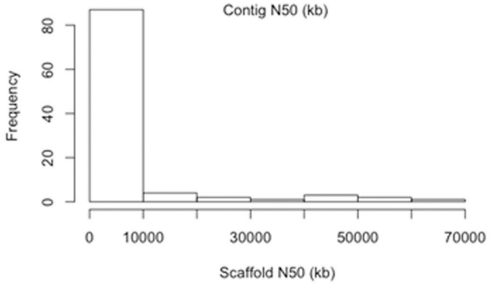
HGAP Preassembly (BLASR):	
Seed length cutoff	16,000 bp
Pre-Assembled Bases	6,281,202,330 bp
Pre-Assembled Reads	464,567
Pre-Assembled N50	18,572 bp

Output (Celera Assembler):	
Number of Polished Contigs	625
Max Contig Length	7,984,151 bp
N50 Contig Length	2,386,328 bp
Sum of Contig Lengths	243,174,629 bp

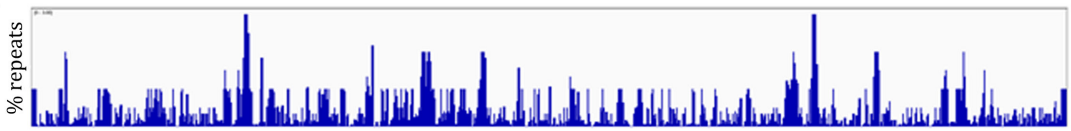
d



e



f



g

h

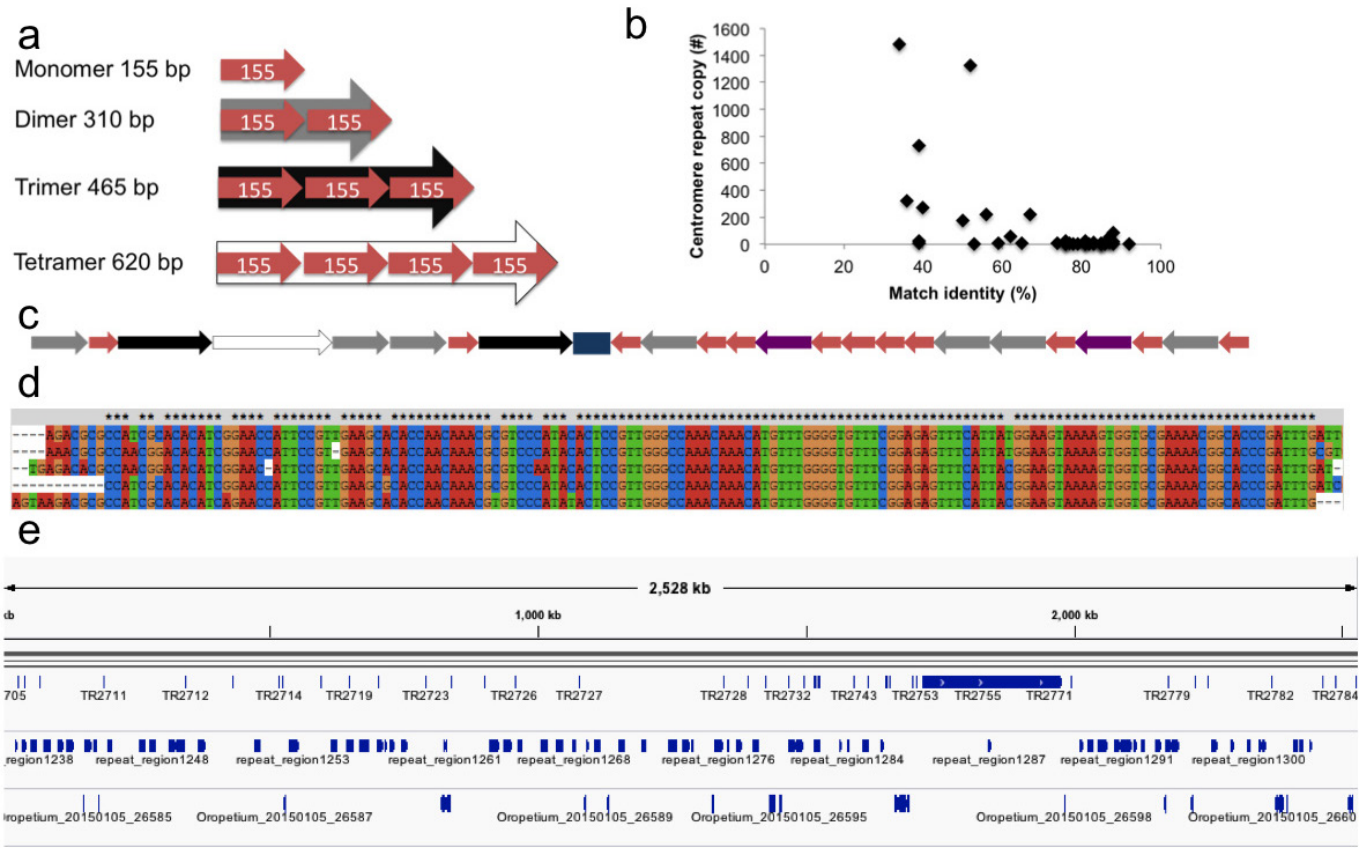
Errors in PacBio assembly	
Number of homozygous SNPs	5,729
Number of homozygous InDels	10,687
Estimated accuracy	99.9999%

Within genome heterozygosity	
Number of heterozygous SNPs	172,250
Number of heterozygous InDels	40,249
Estimated within genome heterozygosity	0.09%

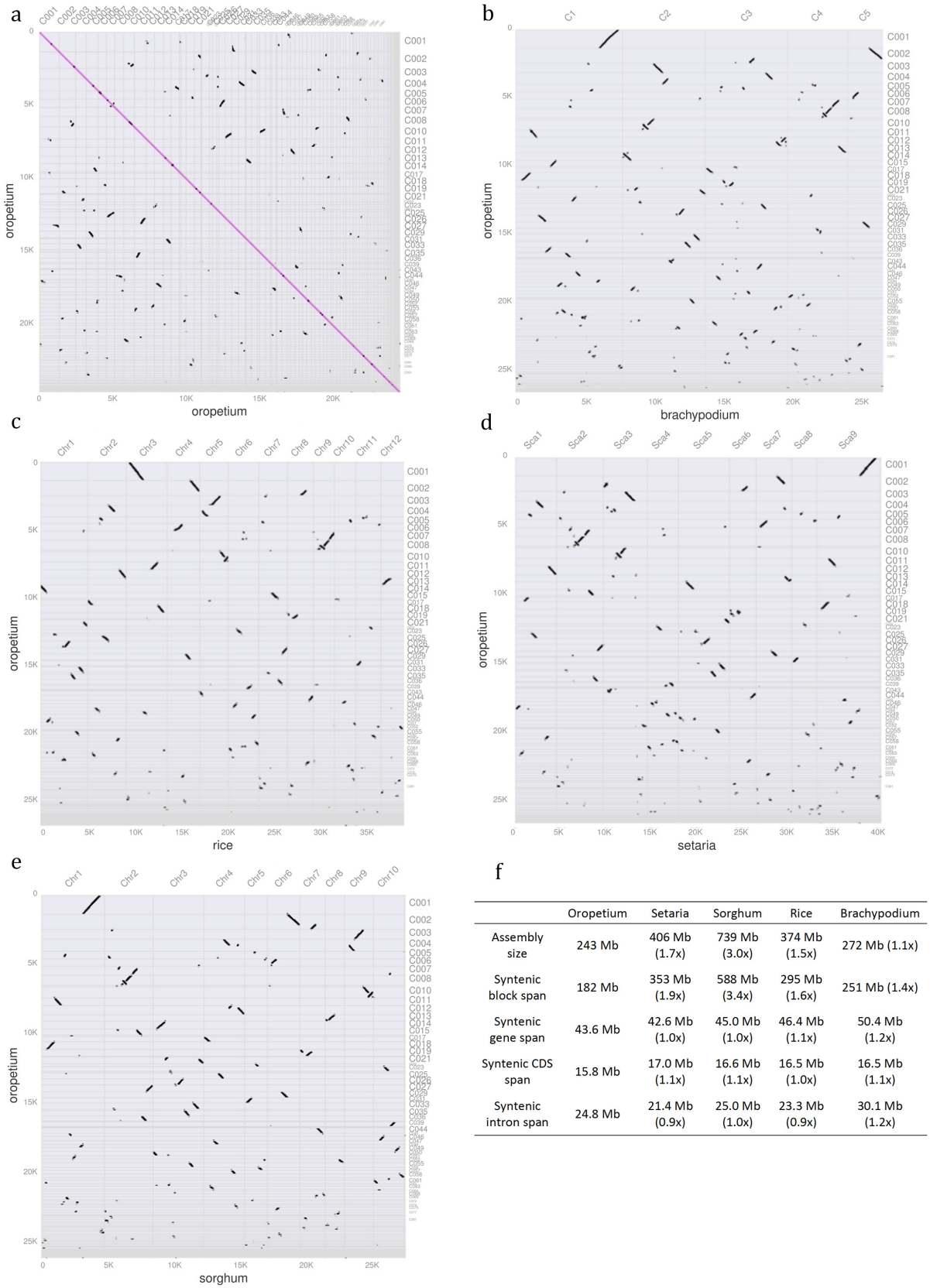
i

	HGAP	Falcon	MHAP
Assembly Size	244Mb	299Mb	240Mb
Contig N50 length	2.4Mb	1.9Mb	1.1Mb
# Contigs	625	612	941
Largest Contig	7.9Mb	7.0Mb	6.3Mb

Extended Data Figure 1 | Summary of the *Oropetium* genome assembly statistics. a, Histogram of length distribution of raw P6C4 chemistry PacBio reads. The mean read length of the raw reads is 12,872 bp, and the N50 is 16,485 bp. **b**, Genome size estimation using *k*-mer distribution. *K*-mer distribution of unassembled *Oropetium* Illumina WGS reads. *K*-mer frequency displays a unimodal curve indicating a low rate of heterozygosity in the *Oropetium* genome. Frequency distribution suggests a genome size of ~245 Mb, consistent with flow-cytometry-based estimations. **c**, SMRT sequencing raw read, preassembly and assembly statistics. **d**, **e**, The distribution of the contig N50 length (**d**) and scaffold N50 length (**e**) of all published plant genomes is plotted. The average contig N50 length for published plant genomes is ~50 kb compared to 2.4 Mb for *Oropetium*. **f**, **g**, Repeat density (as a function of percentage repeats) (**f**) and GC content (**g**) are plotted at a scaled position along each contig. Each contig was divided into 5,000 sliding windows with each window representing 0.02% of the contig length and the averages of each scaled sliding window are plotted. Repeat content and GC content do not vary at the ends of contigs. **h**, Estimated accuracy of SMRT PacBio assembly and within-genome heterozygosity. **i**, Comparison of HGAP Falcon and MHAP PacBio assemblers.



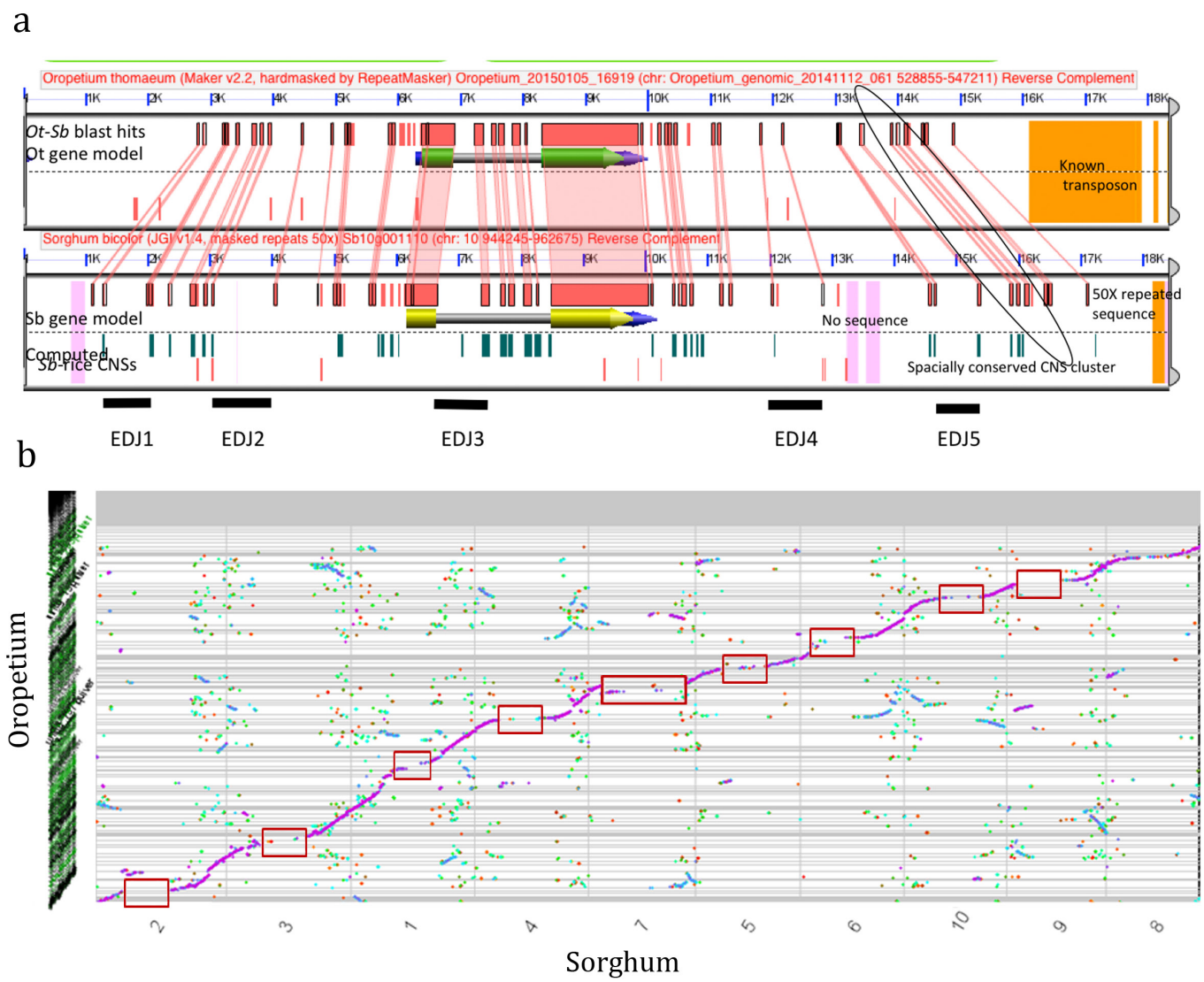
Extended Data Figure 2 | PacBio sequencing and assembly completely resolves the *Oropetium* centromeres. a, The *Oropetium* centromere repeat base is 155 bp (red arrow), whereas they are also found in dimer (310 bp, grey arrow), trimer (465 bp, black arrow) and tetramer (620 bp, white arrow) form. **b**, As the copy number of a repeat increases, the match identity between monomers in the repeat decreases. **c**, The inverted repeat structure of the entire centromere on contig028 with a 60 kb spacer (blue box); arrows are as in **a**. **d**, Consensus 155 bp centromere monomer. **e**, Integrated genome browser view of centromere repeat, LTRs and predicted genes on contig028.



f

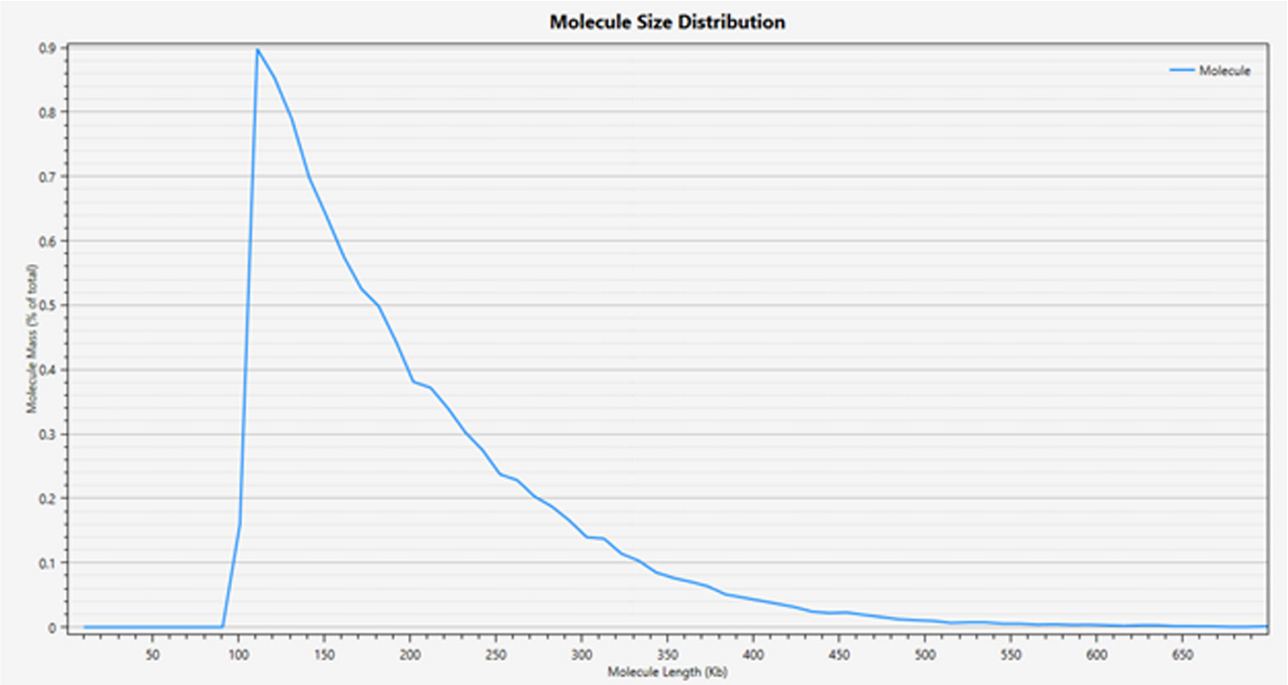
	Oropetium	Setaria	Sorghum	Rice	Brachypodium
Assembly size	243 Mb	406 Mb (1.7x)	739 Mb (3.0x)	374 Mb (1.5x)	272 Mb (1.1x)
Syntenic block span	182 Mb	353 Mb (1.9x)	588 Mb (3.4x)	295 Mb (1.6x)	251 Mb (1.4x)
Syntenic gene span	43.6 Mb	42.6 Mb (1.0x)	45.0 Mb (1.0x)	46.4 Mb (1.1x)	50.4 Mb (1.2x)
Syntenic CDS span	15.8 Mb	17.0 Mb (1.1x)	16.6 Mb (1.1x)	16.5 Mb (1.0x)	16.5 Mb (1.1x)
Syntenic intron span	24.8 Mb	21.4 Mb (0.9x)	25.0 Mb (1.0x)	23.3 Mb (0.9x)	30.1 Mb (1.2x)

Extended Data Figure 3 | Macrosynteny patterns and comparative genomics between the grasses. a–**e**, Macrosynteny of *Oropetium* versus *Oropetium* (**a**); *Oropetium* versus *Brachypodium* (**b**); *Oropetium* versus rice (**c**); *Oropetium* versus *Setaria* (**d**); and *Oropetium* versus *Sorghum* (**e**). **f**, Genome compaction in *Oropetium* compared to related grass genomes. Syntenic block span is based on regions that show conserved synteny across all five genomes. Syntenic gene and coding DNA sequences span is based on 13,683 genes that are retained as genes in orthologous locations across all five genomes. The ratio compared to *Oropetium* is given in brackets.

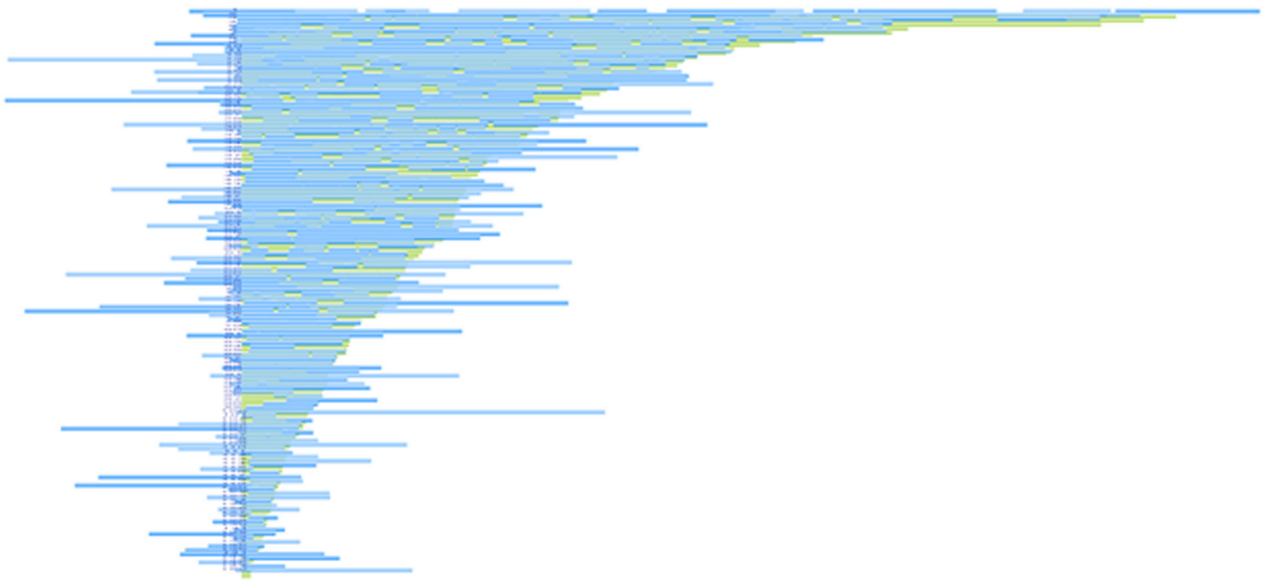


Extended Data Figure 4 | Expansion of intragenic and pericentric regions in *Sorghum* compared to *Oropetium*. **a**, A GEvo sequence similarity graphic of an *Oropetium* gene (upper) and its orthologous *Sorghum* gene (lower). Blast hits (high-scoring segment pairs) are denoted by red rectangles, and syntenic hits are connected by a red line. The green rectangles on the model line of *Sorghum* are conserved noncoding sequences (CNS) computed between *Sorghum* and rice; the expanse of CNS coverage defines 'gene space'. Within the oval are three CNS that may be spatially constrained. The expanded interspersed sequences are annotated at the bottom in black. **b**, Pericentric region expansion in *Sorghum* compared to *Oropetium*. A syntenic dot plot of the *Sorghum* and *Oropetium* genomes is plotted. *Oropetium* contigs are ordered based on synteny with *Sorghum*. Hits are coloured based on K_s divergence, with purple blocks corresponding to 1:1 orthologous regions and other colours corresponding to retained genes from the rho and sigma WGDs. Pericentric regions in *Sorghum* have few syntenic matches to *Oropetium*, suggesting that much of the expansion occurred in pericentric regions.

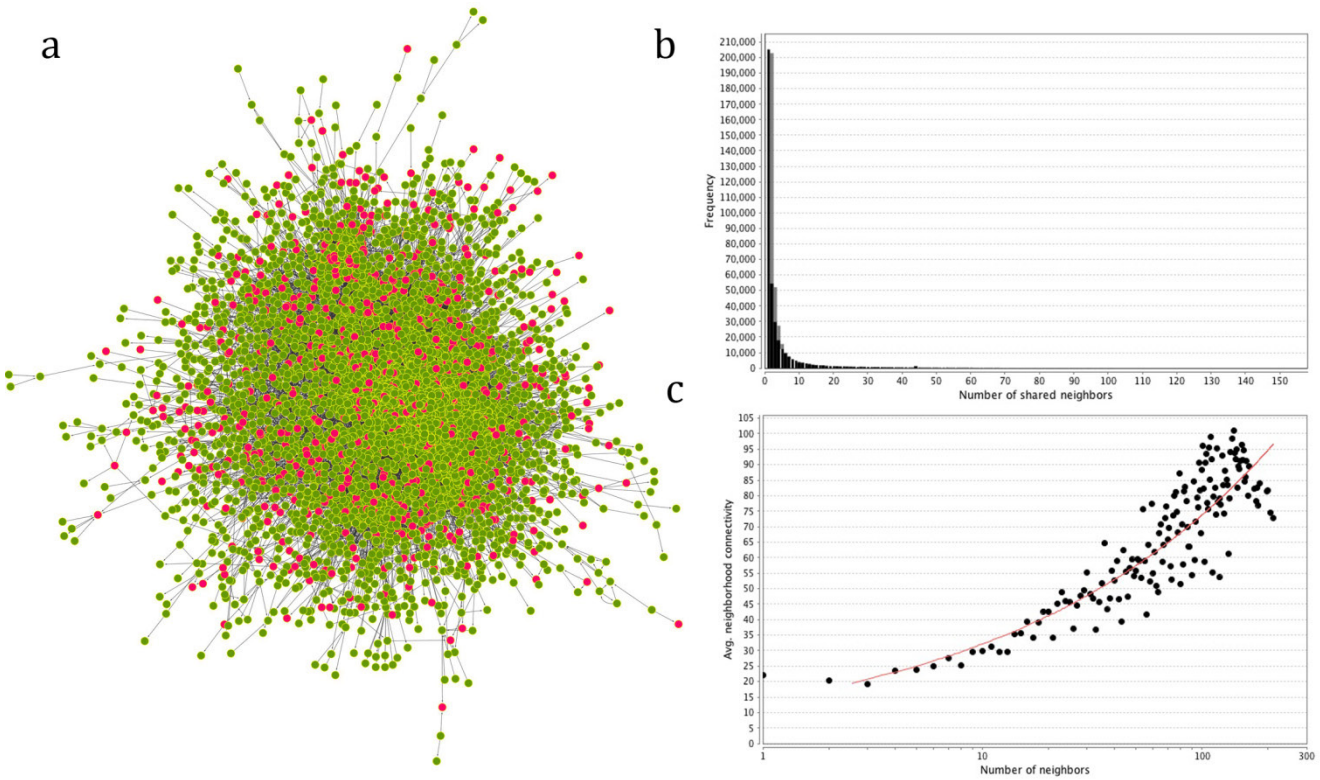
a



b



Extended Data Figure 5 | Assembly improvement using a BioNano-based genome map from the Irys system. a, Distribution of molecule size for raw single-molecule genome mapping data. Size of single molecules in nanochannel arrays is plotted. **b**, Integration of the genome map with the genome assembly. Overlap between the PacBio-based contigs and the genome map. Each line shows a single PacBio contig in green; genome maps are shown in light blue.



Extended Data Figure 6 | Network statistics for tandem duplicated genes. a, Tandem duplicated genes in the metabolic network are shown in pink. **b**, Distribution of shared neighbours. **c**, The average number of neighbours.

Extended Data Table 1 | Telomere repeat (AAACCCT) locations and organization in the *Oropetium* genome

Contig Name	Contig Size (bp)	Start of telomeric array	End of telomeric array	Size of centromeric array (bp)	Teleomeric repeat sequence	Position of Telomere on contig	Number of telomeric repeats
Oropetium_genomic_143	99,304	1	6,446	6,445	AACCCTA	start	910.1
Oropetium_genomic_058	1,564,795	1,560,696	1,564,795	4,099	AGGGTTT	end	580.9
Oropetium_genomic_552	22,498	18,622	22,498	3,876	GTTTAGG	end	562.9
Oropetium_genomic_043	1,920,679	1	3,643	3,642	CCCTAAA	start	515.7
Oropetium_genomic_050	1,822,802	1	3,223	3,222	CCTAAAC	start	453.3
Oropetium_genomic_125	248,855	1	3,182	3,181	AAACCCT	start	452.4
Oropetium_genomic_027	2,706,558	2	2,092	2,090	AAACCCT	start	301.9
Oropetium_genomic_169	56,172	54,243	56,170	1,927	TTAGGGT	end	279.7
Oropetium_genomic_103	526,141	524,277	526,139	1,862	GTTTAGG	end	265.9
Oropetium_genomic_124	262,476	260,617	262,476	1,859	TTTAGGG	end	264
Oropetium_genomic_090	736,395	1	1,601	1,600	CCTAAAC	start	227.4
Oropetium_genomic_010	4,141,579	4,140,107	4,141,579	1,472	TTAGGGT	end	208.7
Oropetium_genomic_076	1,024,162	1	1,169	1,168	CCCTAAA	start	166.1
Oropetium_genomic_493	25,796	24,869	25,795	926	GGGTTTA	end	129.9
Oropetium_genomic_136	153,270	152,446	153,270	824	GTTTAGG	end	119.1
Oropetium_genomic_155	63,826	63,040	63,826	786	TTTAGGG	end	110.4
Oropetium_genomic_019	3,122,409	1	347	346	AAACCCT	start	48
Oropetium_genomic_149	80,145	1	294	293	AAACCCT	start	40.4

Extended Data Table 2 | rRNA tandem array locations and organization in the *Oropetium* genome

Contig Name	Contig Size (bp)	Start of NOR array (bp)	End of NOR array (bp)	Size of NOR tandem array (bp)	Position of NOR on contig	Number of NOR tandem repeats
Oropetium_genomic_182	51,716	1	51,716	51,716	spans contig	5.7
Oropetium_genomic_265	38,885	1	38,885	38,885	spans contig	4.3
Oropetium_genomic_168	56,772	1	56,772	56,772	spans contig	6.3
Oropetium_genomic_192	48,530	1	42,860	42,860	start	4.7
Oropetium_genomic_214	44,298	31,977	44,298	12,321	start	1.3
Oropetium_genomic_539	23,633	1	20,975	20,975	start	2.3

Extended Data Table 3 | Repeat annotation of the *Oropetium* genome

Repeat Class	Number of Elements	Percent Base Pairs Covered
Retrotransposon	214,698	35.60%
Long terminal repeat (LTR)	107,010	25.50%
Gypsy (RLG)	83,872	21.80%
Copia (RLC)	18,223	36.90%
Penelope (RPX)	1,548	0.15%
Unknown LTR (RLX)	3,367	0.44%
LINE (RIL)	17,399	1.90%
SINE (RSX)	2,735	0.07%
DIRS (RYD)	5,098	3.00%
Unknown retrotransposon (RXX)	82,456	7.50%
DNA transposon	69,217	8.50%
Maverick (DMX)	68	0.01%
TIR (DTX)	41,930	6.60%
Unknown DNA transposon (DXX)	27,219	1.90%
No category	7,902	1.00%
Total	**291,817**	**43.80%**

Extended Data Table 4 | Comparisons of repeats and coding features in the monocotyledons

						Transcript statistics				Exon statistics			Intron statistics		
Common name	**Species name**	**chr. #**	**genome size (Mb)**	**repeat #**	**Gene #**	**Avg. Length**	**Median Length**	**Avg. Num Exons**	**Median Num Exons**	**Count**	**Avg. Length**	**Median Length**	**Count**	**Avg. Length**	**Median Length**
Greater duckweed	Spirodela polyrhiza	20	150	23	19,519	4,718	3,015	5.22	3	101,867	222	129	82,368	757	202
Oropedium	Oropetium thomaeum	**9**	**250**	**43**	**28,446**	**2,729**	**1,928**	**4.55**	**3**	**129,421**	**210**	**126**	**100,975**	**446**	**168**
brachy	Brachypodium distachyon	5	272	21	42,868	3,819	3,128	5.38	4	154,738	254	137	120,380	402	142
rice	Oryza sativa	12	403	35	66,338	3,191	2,701	4.4	3	238,247	331	162	177,497	389	166
setaria	Setaria italica	9	510	40	29,448	3,299	2,563	4.96	3	134,802	261	137	106,488	436	145
sorghum	Sorghum bicolor	10	818	62	40,599	2,745	2,189	4.74	3	160,151	252	140	122,497	326	133
corn	Zea mays	10	2,300	85	63,540	4,236	2,747	4.6	3	203,643	238	133	149,177	670	154

LETTER

doi:10.1038/nature15763

Sweet and bitter taste in the brain of awake behaving animals

Yueqing Peng[1,2,3], Sarah Gillis-Smith[1,2,3], Hao Jin[1,2,3], Dimitri Tränkner[1,2,3,4], Nicholas J. P. Ryba[5] & Charles S. Zuker[1,2,3,4]

Taste is responsible for evaluating the nutritious content of food, guiding essential appetitive behaviours, preventing the ingestion of toxic substances, and helping to ensure the maintenance of a healthy diet. Sweet and bitter are two of the most salient sensory percepts for humans and other animals; sweet taste allows the identification of energy-rich nutrients whereas bitter warns against the intake of potentially noxious chemicals[1]. In mammals, information from taste receptor cells in the tongue is transmitted through multiple neural stations to the primary gustatory cortex in the brain[2]. Recent imaging studies have shown that sweet and bitter are represented in the primary gustatory cortex by neurons organized in a spatial map[3,4], with each taste quality encoded by distinct cortical fields[4]. Here we demonstrate that by manipulating the brain fields representing sweet and bitter taste we directly control an animal's internal representation, sensory perception, and behavioural actions. These results substantiate the segregation of taste qualities in the cortex, expose the innate nature of appetitive and aversive taste responses, and illustrate the ability of gustatory cortex to recapitulate complex behaviours in the absence of sensory input.

In mice, sweet and bitter activate cortical fields in the insula (taste cortex) that are separated topographically by approximately 2 mm (ref. 4) (Fig. 1a and Extended Data Fig. 1). We hypothesized that if these cortical fields represent sweet and bitter percepts, their direct activation would evoke 'bitter and sweet sensation' even in the absence of an actual bitter or sweet stimulus. To optogenetically control activation of the gustatory cortex, we introduced channelrhodopsin[5] (ChR2) to the insula of wild-type mice by stereotaxic injection of adeno-associated virus (AAV) targeted to either the bitter or the sweet cortical field (see Fig. 1a, b, Extended Data Fig. 1, Supplementary Table 1 and Methods for details). Single unit recordings of the insular cortex of transduced animals demonstrated that photostimulation evoked reliable neuronal firing that was phase locked to light delivery (Fig. 1c and Extended Data Fig. 1b).

We reasoned that optogenetic activation of the sweet cortical field should trigger behavioural attraction, whereas stimulation of the bitter field should cause strong behavioural avoidance. We used a place-preference test[6] where animals expressing ChR2 in the sweet cortex were introduced to a two-chamber arena in which presence in one of the two chambers was coupled to optogenetic stimulation, in the absence of any reward or punishment; we then determined the animal's preference index as a measure of the time spent in the chamber that was coupled with light stimulation. When the sweet cortical field was stimulated, animals developed strong preference for the chamber coupled to ChR2 stimulation (Fig. 1d and Extended Data Fig. 2). This preference could be transferred to either side of the arena by switching the chamber coupled to the laser stimulation of sweet cortex (Fig. 1d, compare chamber 1 versus chamber 2). When the same sets of experiments were performed in animals expressing ChR2 in the bitter cortical field, mice now displayed a range of unconditioned aversive behaviours (see next section), and after just a few sessions strongly avoided the chamber linked to photostimulation (Fig. 1e). Mice injected with a control AAV expressing enhanced green fluorescent protein (AAV–eGFP construct) exhibited no significant place preference after laser stimulation of either the sweet or bitter cortical fields (Extended Data Fig. 2b). Together, these observations demonstrate that neurons in the sweet and bitter cortical fields drive attractive and aversive responses, respectively.

Next, we examined if activation of the bitter and sweet cortical fields evokes classical taste behaviours[7]. We hypothesized that optogenetic activation of the bitter cortical field should trigger strong light-dependent suppression of licking, while activation of the sweet cortical field should trigger appetitive responses.

We used a behavioural test where motivated animals (thirsty) were trained to lick water in response to a combination visual/tone cue in a head-restrained set-up[8] (see Methods). We then subjected the trained animals expressing ChR2 in the bitter cortical field to testing sessions consisting of a series of water-only trials, but in half of the trials the bitter cortical field was stimulated upon contact of the tongue with the water spout.

During the entire session we imaged (facial features), recorded, and measured licking responses. Figure 2 demonstrates that when the bitter cortical field was stimulated, there was a dramatic suppression of licking behaviour (see also Supplementary Video 1), with the animal's response closely following the ChR2 activation of the bitter cortex. Notably, after strong laser stimulation (10–20 mW), the animals displayed prototypical taste rejection orofacial responses, sometimes including gagging (gaping[9]), and attempts to clean and rid the mouth of the non-existent bitter tastant (Supplementary Video 1; see legend for details).

What about the sweet cortical field? A characteristic feature of sweet taste is that non-thirsty animals remain robustly attracted to sweet solutions, even though they exhibit limited interest for water[10]. Therefore, we predicted that a mildly water-satiated animal expressing ChR2 in the sweet cortical field would still show little attraction for water in control trials (referred to as off-trials), but would exhibit significantly enhanced licking during water trials coupled to laser stimulation of the sweet cortical field (referred to as on-trials). Importantly, the experiment was set up such that the laser shutter was under contact-licking operation, so the animal had control of its own stimulation during the on-trials, and therefore only persistent licking (self-stimulation) would continue to activate the sweet cortex. Our results demonstrate that animals aggressively self-stimulated during on-trial sessions, with ChR2 activation of the sweet cortical field radically increasing licking behaviour, even though the spout still delivered only water, as in the off-trials (Fig. 2b, d).

Just as a lot of sugar can 'mask' a bitter tastant, we hypothesized that strong activation of the cortical field representing sweet taste might be capable of overcoming the natural aversion to an orally applied bitter

[1]Howard Hughes Medical Institute, Columbia College of Physicians and Surgeons, Columbia University, New York, New York 10032, USA. [2]Departments of Biochemistry and Molecular Biophysics, Columbia College of Physicians and Surgeons, Columbia University, New York, New York 10032, USA. [3]Department of Neuroscience, Columbia College of Physicians and Surgeons, Columbia University, New York, New York 10032, USA. [4]HHMI/Janelia Farm Research Campus, 19700 Helix Drive, Ashburn, Virginia 20147, USA. [5]National Institute of Dental and Craniofacial Research, National Institutes of Health, Bethesda, Maryland 20892, USA.

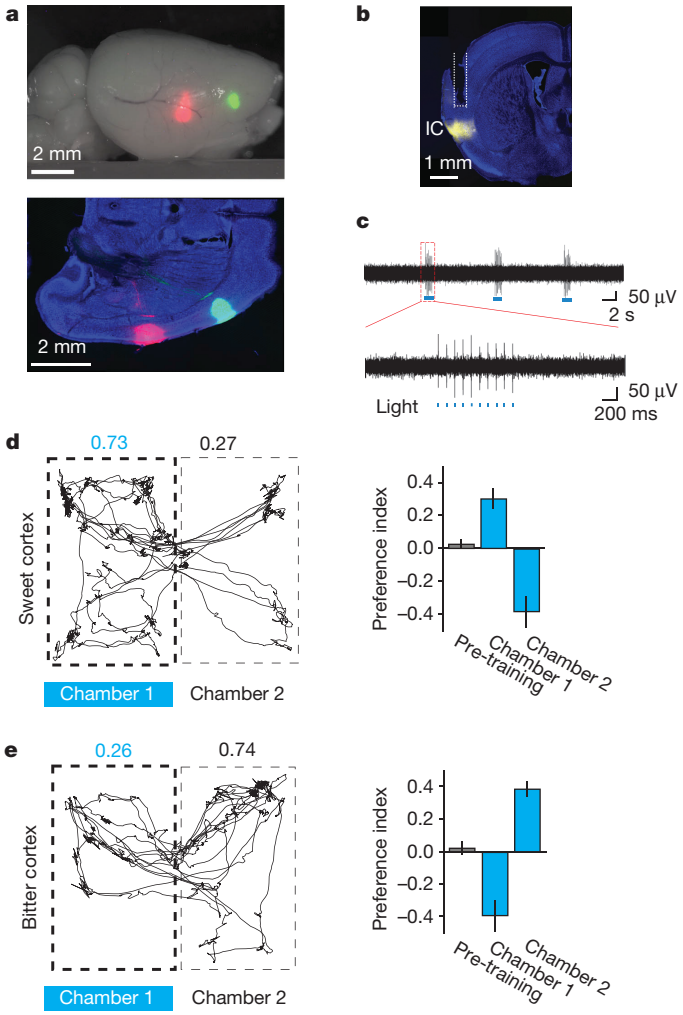


Figure 1 | Place preference by photostimulation of the sweet and bitter cortical fields. a, Sample injection of reporters in stereotactic coordinates defining the sweet and bitter cortical fields. Top: sweet cortex labelled with AAV–GFP and bitter cortex with AAV–TdTomato; bottom: a horizontal section. See Extended Data Fig. 1 for additional data. **b**, Coronal section of a mouse brain (bregma −0.2) stained with TO-PRO-3 (blue). Shown is a representative histological sample of the bitter cortical field expressing ChR2 fused to yellow fluorescent protein (ChR2–YFP), illustrating the location and trajectory (dotted lines) of the implanted guide cannula; IC, insular cortex. **c**, *In vivo* recording of ChR2-expressing insular cortical neurons in response to light stimulation (ten pulses, 10 Hz). The expanded traces show responses to each light pulse (blue bars below the trace). **d**, Left: representative tracking of a mouse during the 5 min preference test in a two-chamber arena; chamber 1 was coupled to light stimulation of the sweet cortical field during the training sessions. Shown are the fractions of time spent in each chamber. Right: quantitation of preference index before (pre-) and after (chamber 1) training with photostimulation of the sweet cortical field ($n = 13$ animals; Mann–Whitney *U*-test, $P < 0.003$). Preference can be readily reversed by light stimulation in the opposing side (chamber 2, $n = 6$; $P < 0.02$). **e**, Representative mouse track and quantitation of preference index in mice expressing ChR2 in the bitter cortical field; note significant aversion to the chamber coupled to photostimulation (chamber 1, $n = 15$; Mann–Whitney *U*-test, $P < 0.005$); this behavioural aversion can be switched to the opposite chamber by re-exposure to photostimulation in chamber 2 ($n = 4$; $P < 0.03$). Values are mean ± s.e.m. See Extended Data Fig. 2b for GFP control injections.

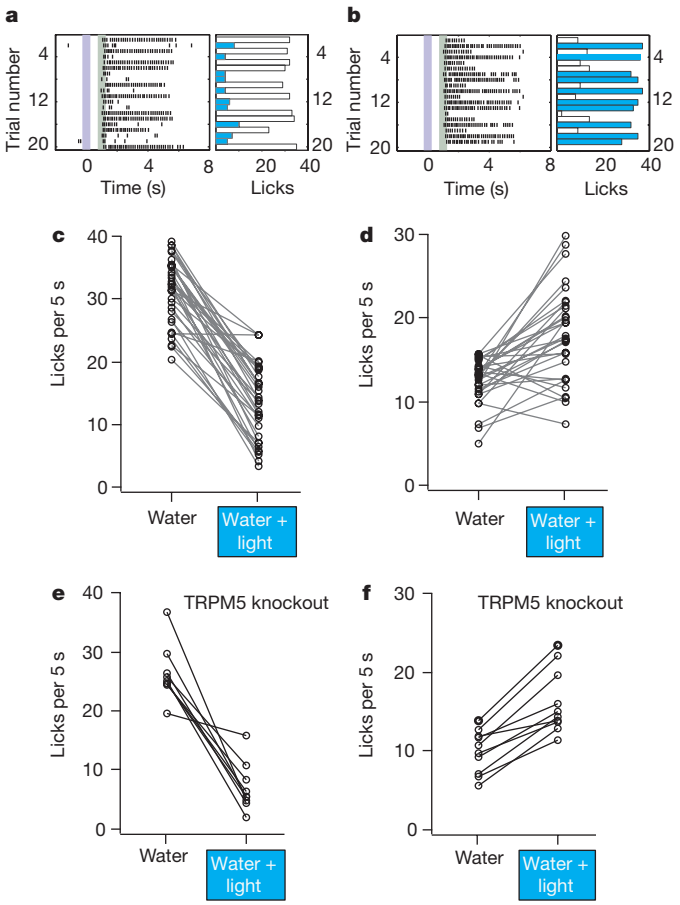


Figure 2 | Photostimulation of bitter and sweet cortical fields drives aversive and appetitive behaviours. a, **b**, Representative raster plots (left) and histograms (right) illustrating licking events during a 5 s licking window in the presence (blue) or absence (open) of light stimulation of (**a**) the bitter and (**b**) the sweet cortical fields. The purple line at time zero indicates the start of each trial; the green line indicates the onset of water delivery. **c**, **d**, Quantitation of licking responses with and without light stimulation in (**c**) the bitter cortical field ($n = 34$, Mann–Whitney *U*-test, $P < 4 \times 10^{-12}$) or (**d**) sweet cortical field ($n = 31$, Mann–Whitney *U*-test, $P < 5 \times 10^{-5}$) of wild-type mice. **e**, **f**, Quantitation of licking responses in TRPM5 knockout mice (**e**, bitter cortical fields, $n = 9$, Mann–Whitney *U*-test, $P < 5 \times 10^{-5}$; **f**, sweet cortical fields, $n = 10$, Mann–Whitney *U*-test, $P = 0.001$). Each point indicates data from an individual mouse before and after photostimulation.

stimulus. Therefore, we asked whether photostimulation of the sweet cortical field in animals expressing ChR2 in sweet cortex could switch preference for an otherwise aversive tastant. Conversely, we also tested whether photostimulation of the bitter cortical field triggers aversion to an otherwise sweet, attractive tasting chemical. Our results (Extended Data Fig. 3) show both postulates to be correct, and highlight how activation of selective taste cortical fields can mask the hedonic value of oral taste stimulation.

The experiments described above show that direct control of primary taste cortex can evoke specific, reliable, and robust behaviours naturally symbolic of taste responses to chemical tastants. These gain-of-function studies also illustrate how top-down control of the taste pathway can activate innate, immediate responses representing sweet and bitter taste.

To formally demonstrate that these cortically triggered behaviours are innate (that is, independent of learning or experience) we performed similar stimulation experiments in mice that had never tasted sweet or bitter chemicals (TRPM5 null mice[10]; Extended Data Fig. 4). Indeed, our results (Fig. 2e, f) showed that even in animals that had never experienced sweet or bitter taste, ChR2 activation of the corresponding cortical fields still triggered the appropriate behavioural response, thus substantiating the predetermined nature of the sense of taste.

It has been known for a long time that decerebrated animals can still exhibit stereotyped attraction and aversion to sweet and bitter chemicals[11]. This is thought to be mediated by brainstem taste circuits dedicated to immediate responses[11,12]. Therefore, to evaluate the necessity (and sufficiency, see next section) of taste cortex in taste

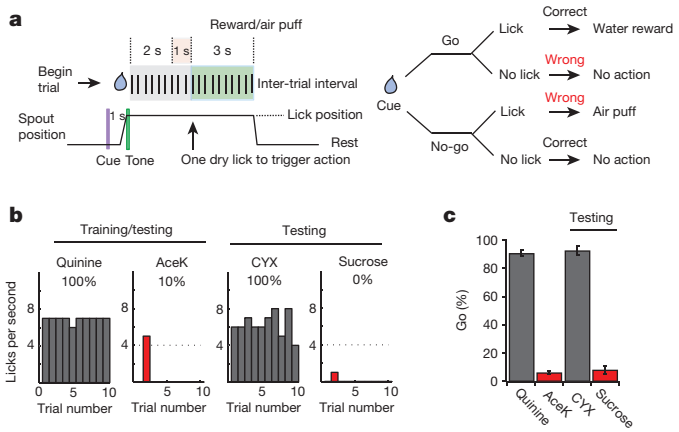


Figure 3 | Go/no-go taste discrimination task in head-restrained mice. **a**, Schematic and flow chart of the go/no-go taste discrimination task. Each trial starts with a visual cue (purple line), followed 1 s later by a tone (green line) to alert mice to initiate licking. After sampling, mice were given 3 s to continue to lick (go) or withhold licking (no-go) in response to the test tastant. For go trials, mice were rewarded with water (3 s) if they chose to lick within the 3-s interval. For no-go trials, mice received a mild air puff to the eyelid if they failed to withhold licking. After the reward/penalty phase, the spout retracted and was cleared for the next trial; inter-trial intervals were 8 s. **b**, Representative histograms illustrating recognition and generalization within bitters and sweets. This animal was trained and tested with 4 mM AceK (sweet no-go) and 0.5 mM quinine (bitter go), and then assayed with 100 mM sucrose and 10 μM cycloheximide (CYX). **c**, Quantitation in nine animals, demonstrating highly reliable taste recognition and discrimination. Values are mean ± s.e.m.

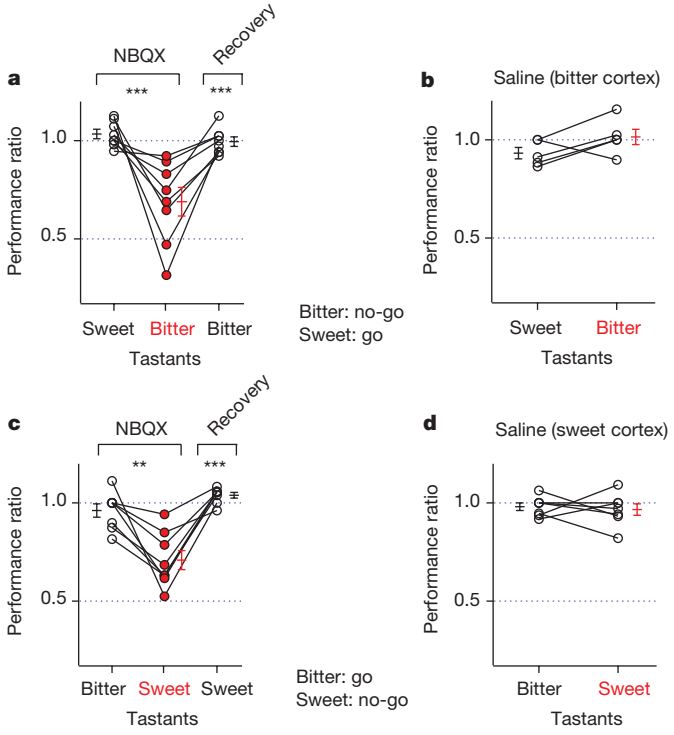


Figure 4 | Inactivation of the bitter and sweet cortical fields disrupts taste discrimination. **a**, Quantitation of performance ratios (see methods) before and after bilateral silencing (NBQX, 5 mg ml^{-1}) of the bitter cortical fields ($n = 8$); animals were trained to no-go to bitter and go to sweet. Note the impact in bitter taste discrimination, but no significant effect in sweet taste. After washout of the drug, the animal's ability to recognize bitter is restored. Comparable results are obtained when animals are instead trained to go to bitter and no-go to sweet (Extended Data Fig. 2). **b**, Quantitation of performance ratios with saline controls in bitter cortical fields; there is no significant effect on sweet or bitter taste ($n = 5$; Mann–Whitney U-test, $P = 0.14$). **c**, Quantitation of performance ratios with bilateral injection of NBQX in the sweet cortical fields ($n = 8$). Animals were trained to no-go to sweet and go to bitter; note significant deficit in sweet taste, but no effect on bitter taste. After washout of the drug, the animal's ability to recognize sweet is restored. **d**, Saline injections in the sweet cortical fields have no significant effect on bitter or sweet taste ($n = 7$; Mann–Whitney U-test, $P = 0.80$). Values are mean ± s.e.m. Mann–Whitney U-test, $^{**}P < 0.01$, $^{***}P < 0.001$.

recognition and discrimination, we needed to design a test that bypasses immediate taste responses, and instead engages cortical circuits. In this assay (go or no-go behavioural test)[13,14], thirsty animals were trained to sample a test tastant from a spout, and then to report its identity either by licking (go) or withholding licking (no-go) (Fig. 3). This learned behaviour required the animal to sample the cue, recognize the tastant, and execute the appropriate behaviour in each trial. We trained animals several ways, including to go to bitter and no-go to sweet, exactly the opposite of the innate drive. After 10–15 sessions of training (each consisting of 80 trials, with 40 randomly presented sweet and 40 bitter cues), mice were able to report the tastant's identity with almost 90% accuracy (Fig. 3). To further demonstrate the selectivity of the assay and responses, we next tested the animals with sweet and bitter chemicals not used in the training phase. Given that all sweet tastants activate the same sweet taste receptor[15–17], and all bitters the same class of taste receptor cells[18], we expected that novel sweets should also be recognized as no-go cues, whereas novel bitters should be seen as go cues. Indeed, animals trained with the bitter tastant quinine and the artificial sweetener acesulfame K (AceK) recognized and responded with similar accuracy to cycloheximide and sucrose, bitter and sweet tastants with completely different chemical structures from the training set (Fig. 3c).

We implanted cannulae bilaterally into the bitter cortical fields of trained animals (Supplementary Table 1), waited 2 weeks for recovery, and assayed tastant discrimination in the go/no-go behavioural test before and after bilateral injection of a glutamate receptor antagonist (NBQX) to silence cortical activity[19,20]. As shown in Fig. 4, silencing the bitter cortical fields prevented animals from reliably identifying the bitter tastant (see Extended Data Fig. 5 for additional examples using the reverse training test). In contrast, their ability to recognize sweet tastants remained unimpaired. Importantly, the loss of bitter taste function was fully reversible upon washout of the drug (Fig. 4a), whereas injection of a saline control in the bitter cortical fields had no significant effect on either bitter or sweet taste sensing (Fig. 4b). We used the same strategy to conduct loss-of-function experiments in the sweet cortex. Indeed, bilateral silencing of the sweet cortical fields disrupted sweet, but not bitter, taste discrimination (Fig. 4c, d). As expected, animals recovered sweet taste perception after drug washout. Taken together, these results substantiate the essential role of the sweet and bitter cortical fields in sweet and bitter taste recognition.

What is the mouse sensing upon direct activation of a taste cortical field? Does optogenetic stimulation create internal representations that mimic those evoked by sweet and bitter chemicals on the tongue? If so, we reasoned that animals trained to recognize and report the sensory features of an orally provided sweet or bitter tastant (for example, in a go/no-go assay) should respond similarly to optogenetic stimulation of the corresponding cortical fields, even though the animal had never been trained with light stimulation. In essence, if light and the chemical tastant evoke similar percepts, then light will generalize to the learned responses associated with the orally supplied stimulus.

We first focused on sweet, because activation of the bitter cortical field evokes prototypical and highly salient orofacial responses that are already strongly indicative of bitter perception (Supplementary Video 1). We introduced ChR2 into the sweet cortical field of untrained mice and validated robust light-triggered appetitive responses (see Fig. 2). Then, the mice were trained in a go/no-go behavioural test where they learned to associate go with a bitter chemical and a low-salt solution (Fig. 5a), and no-go with sweet taste. Critically, under this test, mice needed to report both an aversive (bitter) and an attractive cue (low salt, see also Extended Data Fig. 6) in the same arm of the behavioural

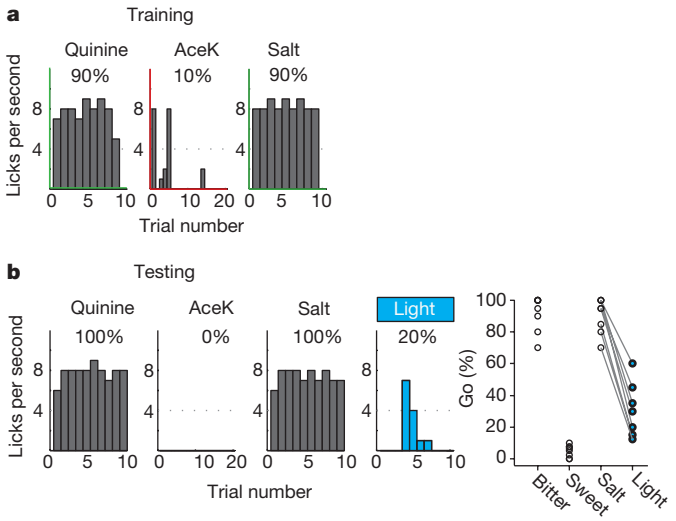


Figure 5 | Cross-generalization between orally supplied taste stimuli and photostimulation of the sweet cortex. a, Representative histograms illustrating mouse performance during a training session in the go/no-go discrimination task. The mouse was trained to go to bitter (0.5 mM quinine) and low salt (20 mM NaCl), and no-go to sweet (4 mM AceK). Note that both bitter (aversive) and low salt (attractive) were used in the same branch of the behavioural task (go) to exclude the valence as an identifier. **b**, Left: representative histograms illustrating cross-generalization between taste stimulation and photostimulation of the sweet cortical field. Right: quantitation of the responses from individual animals to quinine, AceK, salt and salt + light ($n = 8$, Mann–Whitney U-test, $P < 0.0002$).

test, hence removing pure valence[21] as a way of identifying tastants. After mice performed at or above 80% accuracy (Fig. 5a), we assayed whether light (previously triggering strong appetitive responses) was being sensed and reported as sweet (now a no-go response). Animals were tested with 50 randomized trials consisting of 20 bitter, 10 sweet, 10 low salt, and 10 low salt linked to light stimulation of the sweet cortical field. Our results (Fig. 5b) showed that light stimulation of sweet cortex was indeed being sensed as a 'fictive' sweet stimulus, eliciting strong and reliable no-go responses; Extended Data Fig. 7 shows similar experiments and equivalent findings with bitter cortex. Taken together, these results show that activation of a taste cortical field recapitulates an internal representation (for example, perceptual quality) naturally indicative of the orally presented chemical.

The essential role of the sense of taste is to evaluate the quality of a food source or a meal, and to activate the appropriate behavioural actions to consume or reject ingestion[1]. The taste cortex is thought to represent the basic sensory features of the different taste qualities[22,23], and to function as a central neural 'hub' that informs and integrates with other brain areas, and the internal state, to guide taste-dependent actions.

This work centred on the study of the two most distinctive taste qualities, sweet and bitter. These two differ not only in quality but also in valence, mediating innately attractive and aversive behaviours. Many studies have used optogenetics to activate ensembles of neurons and examine their physiological and behavioural consequences[6,24–27]. In this work we explored the internal representation of arguably the two most recognizable chemosensory percepts. Our current studies demonstrate that it is possible to govern an animal's perception and behavioural responses by direct manipulation of selective taste cortical fields. Notably, unlike our other fundamental chemical sense (smell), activation of the sweet and bitter cortical fields evokes predetermined behavioural programs, independent of learning and experience, further illustrating the hardwired and innate nature of the sense of taste.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 7 April; accepted 29 September 2015.
Published online 18 November 2015.

1. Lindemann, B. Receptors and transduction in taste. *Nature* **413,** 219–225 (2001).
2. Yamamoto, T. Taste responses of cortical neurons. *Prog. Neurobiol.* **23,** 273–315 (1984).
3. Accolla, R., Bathellier, B., Petersen, C. C. & Carleton, A. Differential spatial representation of taste modalities in the rat gustatory cortex. *J. Neurosci.* **27,** 1396–1404 (2007).
4. Chen, X., Gabitto, M., Peng, Y., Ryba, N. J. & Zuker, C. S. A gustotopic map of taste qualities in the mammalian brain. *Science* **333,** 1262–1266 (2011).
5. Boyden, E. S., Zhang, F., Bamberg, E., Nagel, G. & Deisseroth, K. Millisecond-timescale, genetically targeted optical control of neural activity. *Nature Neurosci.* **8,** 1263–1268 (2005).
6. Lammel, S. *et al.* Input-specific control of reward and aversion in the ventral tegmental area. *Nature* **491,** 212–217 (2012).
7. Halpern, B. P. in *Drinking Behavior* (eds Weijnen, J. A. W. M. & Mendelson, J.) 1–92 (Springer, 1977).
8. Guo, Z. V. *et al.* Procedures for behavioral experiments in head-fixed mice. *PLoS ONE* **9,** e88678 (2014).
9. Grill, H. J. & Norgren, R. The taste reactivity test. I. Mimetic responses to gustatory stimuli in neurologically normal rats. *Brain Res.* **143,** 263–279 (1978).
10. Zhang, Y. *et al.* Coding of sweet, bitter, and umami tastes: different receptor cells sharing similar signaling pathways. *Cell* **112,** 293–301 (2003).
11. Grill, H. J. & Norgren, R. The taste reactivity test. II. Mimetic responses to gustatory stimuli in chronic thalamic and chronic decerebrate rats. *Brain Res.* **143,** 281–297 (1978).
12. Reilly, S. & Pritchard, T. C. Gustatory thalamus lesions in the rat: I. Innate taste preferences and aversions. *Behav. Neurosci.* **110,** 737–745 (1996).
13. Gardner, M. P. & Fontanini, A. Encoding and tracking of outcome-specific expectancy in the gustatory cortex of alert rats. *J. Neurosci.* **34,** 13000–13017 (2014).
14. Graham, D. M., Sun, C. & Hill, D. L. Temporal signatures of taste quality driven by active sensing. *J. Neurosci.* **34,** 7398–7411 (2014).
15. Li, X. *et al.* Human receptors for sweet and umami taste. *Proc. Natl Acad. Sci. USA* **99,** 4692–4696 (2002).
16. Nelson, G. *et al.* Mammalian sweet taste receptors. *Cell* **106,** 381–390 (2001).
17. Zhao, G. Q. *et al.* The receptors for mammalian sweet and umami taste. *Cell* **115,** 255–266 (2003).
18. Mueller, K. L. *et al.* The receptors and coding logic for bitter taste. *Nature* **434,** 225–229 (2005).
19. Calu, D. J., Roesch, M. R., Haney, R. Z., Holland, P. C. & Schoenbaum, G. Neural correlates of variations in event processing during learning in central nucleus of amygdala. *Neuron* **68,** 991–1001 (2010).
20. Tye, K. M. *et al.* Amygdala circuitry mediating reversible and bidirectional control of anxiety. *Nature* **471,** 358–362 (2011).
21. Small, D. M. *et al.* Dissociation of neural representation of intensity and affective valuation in human gustation. *Neuron* **39,** 701–711 (2003).
22. Spector, A. C. & Travers, S. P. The representation of taste quality in the mammalian nervous system. *Behav. Cogn. Neurosci. Rev.* **4,** 143–191 (2005).
23. Simon, S. A., de Araujo, I. E., Gutierrez, R. & Nicolelis, M. A. The neural mechanisms of gustation: a distributed processing code. *Nature Rev. Neurosci.* **7,** 890–901 (2006).
24. Witten, I. B. *et al.* Cholinergic interneurons control local circuit activity and cocaine conditioning. *Science* **330,** 1677–1681 (2010).
25. Choi, G. B. *et al.* Driving opposing behaviors with ensembles of piriform neurons. *Cell* **146,** 1004–1015 (2011).
26. Atasoy, D., Betley, J. N., Su, H. H. & Sternson, S. M. Deconstruction of a neural circuit for hunger. *Nature* **488,** 172–177 (2012).
27. Nieh, E. H. *et al.* Decoding neural circuits that control compulsive sucrose seeking. *Cell* **160,** 528–541 (2015).

Supplementary Information is available in the online version of the paper.

Acknowledgements We particularly thank H. Fischman and R. Lessard for suggestions, and members of the Zuker laboratory for comments. We also thank D. Salzman, K. Scott, and R. Axel for discussions. This research was supported in part by a grant from the National Institute of Drug Abuse (DA035025) to C.S.Z., and the Intramural Research Program of the National Institutes of Health, National Institute of Dental and Craniofacial Research (to N.J.P.R.). C.S.Z. is an investigator of the Howard Hughes Medical Institute and a Senior Fellow at Janelia Farms Research Campus, Howard Hughes Medical Institute.

Author Contributions Y.P. designed the study, performed experiments, and analysed data; S.G.-S. performed animals studies, viral injections, histology and analysed data; H.J. performed c-Fos expression studies; D.T. developed the initial behavioural platforms; N.J.P.R. and C.S.Z. designed the study, analysed data, and together with Y.P. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to C.S.Z. (cz2195@cumc.columbia.edu) or N.J.P.R. (nick.ryba@nih.gov).

METHODS

Stereotaxic injections and anatomy. All procedures were performed according to the approved protocols at Columbia University. Six- to eight-week-old C57BL6/J and *Trpm5*$^{-/-}$ mice were used for viral injections. All surgeries were performed using aseptic technique. Mice were anaesthetized with ketamine and xylazine (100 mg/kg body weight and 10 mg/kg body weight, intraperitoneal), placed into a stereotaxic frame, and unilaterally injected with ~30 nl AAV carrying ChR2 (AAV9.CamKIIa.hChR2(H134R)-EYFP.WPRE.SV40, Penn Vector Core) either in the sweet cortical field (bregma 1.6 mm; lateral 3.1 mm; ventral 1.8 mm), or the bitter cortical field (bregma −0.3 mm; lateral 4.2 mm; ventral 2.8 mm). After viral injection, a guide cannula (26 gauge, PlasticsOne) or a customized implantable fibre (200 μm, numerical aperture = 0.39) was implanted 300–500 μm above the injection site, and fixed in place with dental cement. A metal head-post was also attached and secured with dental cement for the purpose of head fixation during behavioural experiments. For pharmacological experiments, AAV–ChR2 was injected bilaterally in the sweet or bitter cortical fields, followed by bilateral implantation of guide cannulae. Mice were allowed to recover for 2–3 weeks before the start of behavioural experiments. Placements of viral injections, guide cannulae, and implantable fibres were histologically verified at the termination of the experiments by TO-PRO3 (1:1,000, Invitrogen) staining of coronal sections (100 μm). Fluorescent images were acquired using a confocal microscope (FV1000, Olympus).

Animals. All behavioural experiments with wild-type animals used 6- to 8-week-old male C57BL6/J mice. No statistical methods were used to predetermine sample size, and investigators were not blinded to group allocation. No method of randomization was used to determine how animals were allocated to experimental groups.

***In vivo* recordings.** Mice expressing ChR2 in taste cortex were anaesthetized with urethane (1.8 mg/g body weight), and the insular cortex was exposed as previously described[4]. Extracellular neural activity was recorded using a tungsten electrode (resistance 2.0–4.0 MΩ, FHC). Data were acquired, amplified, digitized, and bandpass filtered at 600–6,000 Hz with a Neuralynx data acquisition system. For photostimulation, 10 Hz, 5-ms pulses of 473 nm light (~5 mW) were delivered via a solid-state laser (Shanghai Laser & Optics Century Co.) coupled to an optical fibre (200 μm) positioned above the insular cortex.

c-Fos induction and Immunohistochemistry. Individual mice were implanted with an intraoral cannula[28] 3 days before c-Fos induction. On the day of experiments, mice were anaesthetized with urethane (1.6 mg/g body weight) and the trachea was cannulated to aid breathing during oral stimulus presentation. Tastants were perfused into the mouth through the intraoral cannula for 1.5 h at a rate of ~6 ml h^{-1}. Mice were allowed to rest for 30 min and processed for immunostaining as previously described. The brains were sectioned coronally at 100 μm, and labelled with goat anti-c-Fos (Santa Cruz, sc-52-G) overnight; Alexa 488 donkey anti-goat or cy3 donkey anti-goat (Jackson immunoResearch) were used to visualize c-Fos expression. All images were taken using an Olympus FluoView 1000 confocal microscope.

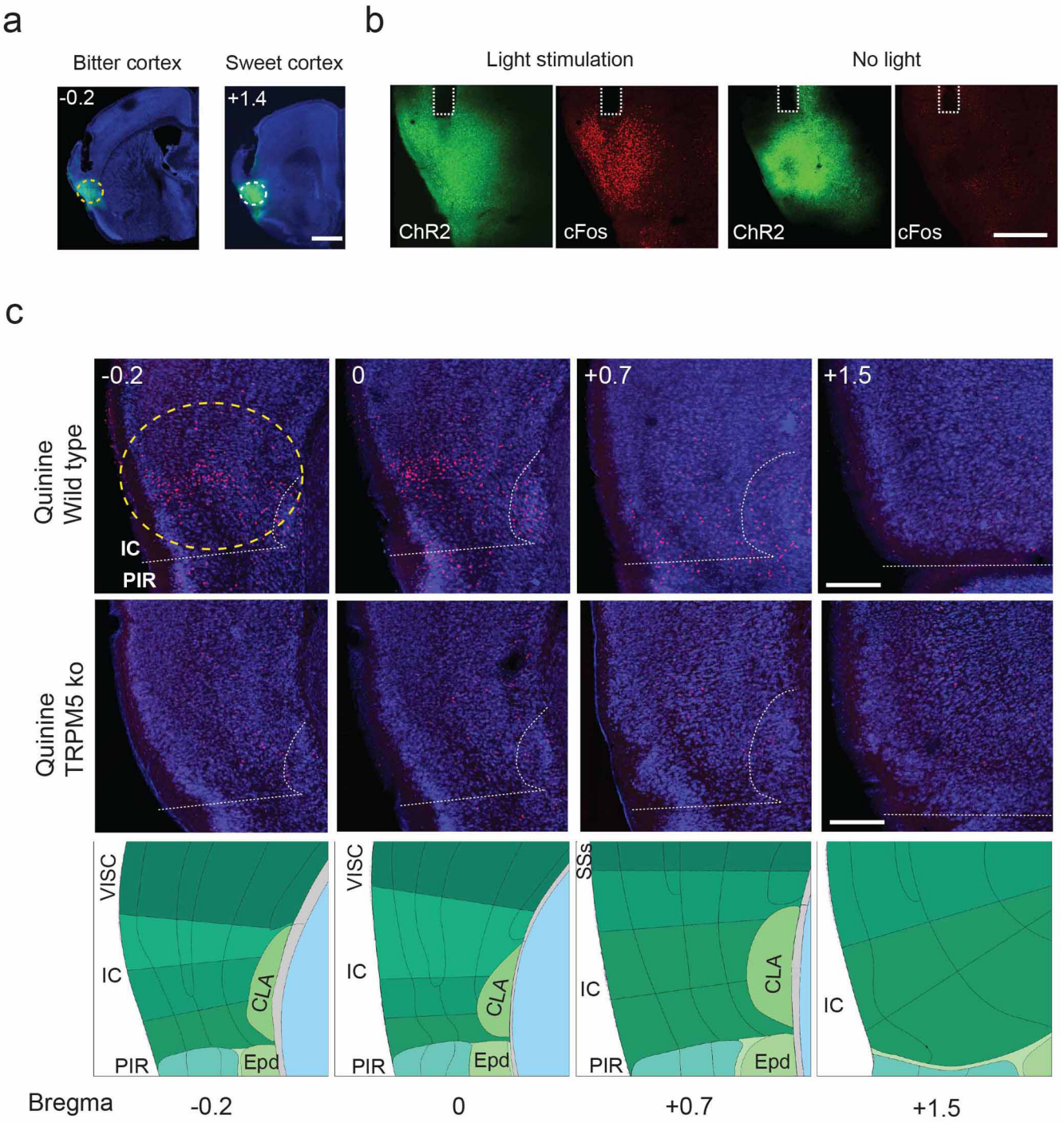
Place preference assays. Individual mice were tested in a custom-built two-chamber arena (30 cm × 30 cm total size). To differentiate the chambers, one chamber was designed with alternating black and white vertical stripes on its walls, whereas the other chamber was uniformly black. The arena was contained within a sound-attenuating cubicle (Med Associates). Mice were trained in the arena for 30 min with photostimulation of the sweet or bitter cortical field, and tested in the absence of any light stimulation for 5 min at the end of each session (defined as 'preference test'). Animal locations were tracked in real time by video imaging. At the beginning of the experiments, mice were acclimated to the arena for one session without light stimulation (defined as the pre-test condition). Photostimulation sessions began the next day, with two daily sessions for about 1 week. For each mouse, one chamber was randomly selected for photostimulation (chamber 1); when a mouse was located in this chamber, light was delivered (20 Hz, 20-ms pulses, 5–10 mW) for 5-s intervals, with 5-s rest periods to avoid over-stimulation or phototoxicity. After 1 week of sessions, a 'reverse probe' study was performed in a subset of animals, during which photostimulation was delivered in the opposing chamber (chamber 2). Animals were trained for a minimum of eight sessions, and the preference tests from the last three sessions were used to calculate the preference index (PI); $PI = (t_1 - t_2)/(t_1 + t_2)$, where t_1 is the fractional time a mouse spent in the chamber 1, and t_2 is the time spent in chamber 2.

Lick preference assays. Mice were first water-deprived for 24 h to motivate drinking behaviour. They were then introduced to head restraint and acclimated to drinking from a motor-positioned spout in 60-trial sessions (15 min), twice a day for 3 days. Each trial began with a flash, followed 1 s later by the spout swinging into position and a tone (4 kHz) to indicate the onset of water delivery. The spout remained in position for 5 s and was then removed. Mice were weighed daily during the habituation period as well as during any behavioural tests requiring water restriction. Additional water was supplied as necessary to ensure that animals maintained at least 85% of their initial body weight. To measure attractive/appetitive responses, mice were mildly water restrained (exhibiting an average of not more than 15 licks per 5-s trial in the lick preference assay), and supplied with approximately 5 μl water during each trial. To measure aversion, mice were water-deprived for 24 h, and supplied with approximately 10 μl water distributed over the full 5 s of spout presentation for each trial (so that animals remained eager to lick for all 5 s). To ensure animals were appropriately motivated in the lick preference behavioural assays (that is, thirsty to examine lick suppression, and mildly satiated to examine attraction), we examined animals exhibiting an average of at least 20 licks per 5 s trial as an indicator of 'thirst', and not more than 15 licks per 5 s trial for mild satiation. Animals were recorded by video for the entire session, and licks were analysed and counted by custom-written MATLAB software (Mathworks). Light stimulation and water delivery were controlled by the same software via an Arduino board. All animals analysed in these studies had histologically confirmed expression of ChR2 in the sweet or bitter cortical fields (Supplementary Table 1).

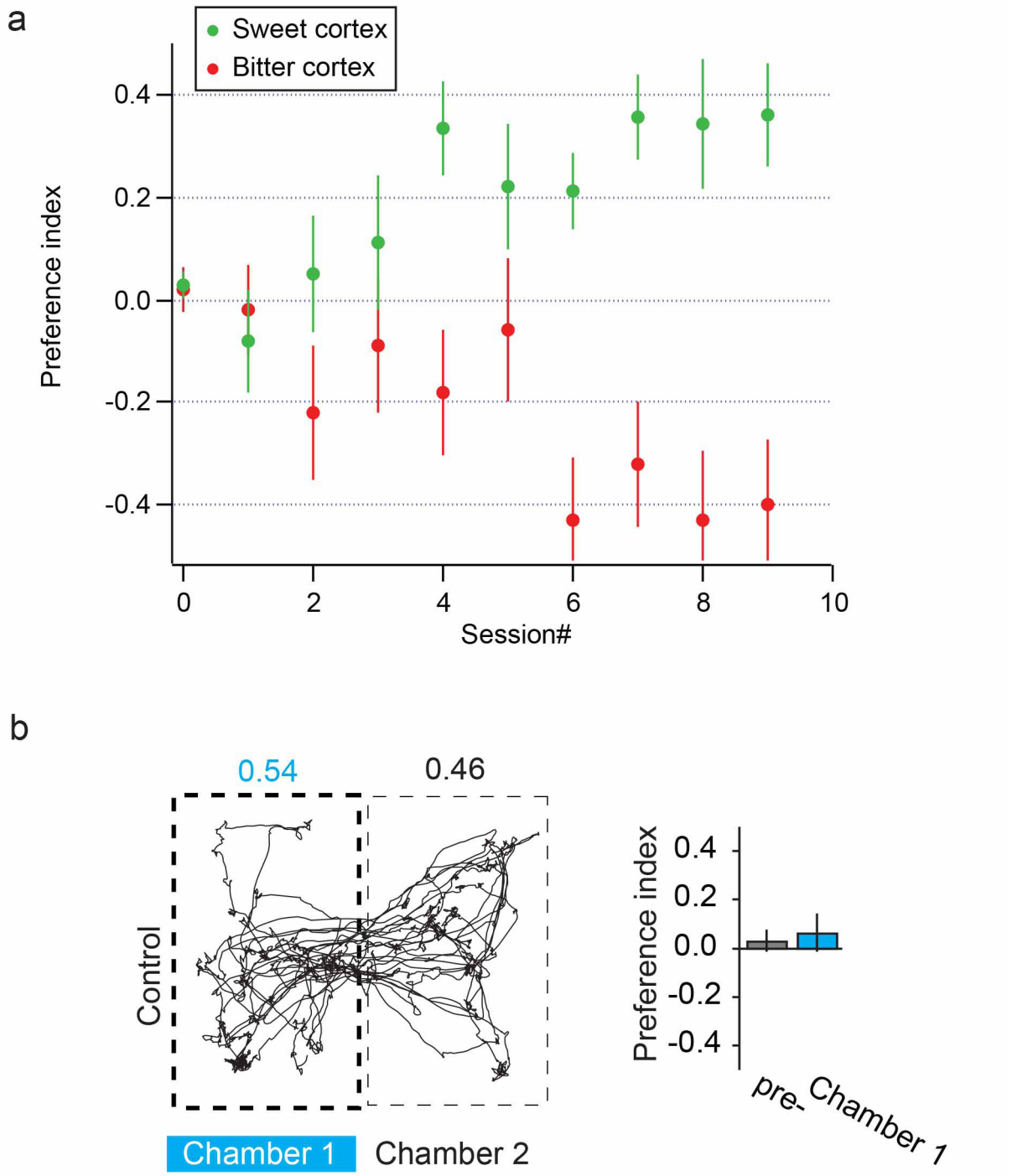
Go/no-go taste discrimination behaviour. Mice deprived of water for 24 h were first acclimated to consuming water in a head-restrained position for 15-min sessions over 2–3 days. Animals were then trained to perform a taste discrimination task, in which they were to lick, and receive a water reward, in response to a 2-μl presentation of tastant-1 ('go') and to withhold licking in response to tastant-2 ('no-go'). The presentation of the go and no-go stimuli was randomized. Each trial began with a visual cue (100-ms light flash), followed 1 s later by a tone (4 kHz, 300 ms) alerting the animal to sample the test tastant (for example, AceK or quinine; ~2 μl per sample). After sampling, mice were given 3 s either to continue to lick the spout (go trial) or to withhold licking (no-go trials). On go trials, if a mouse chose to lick within the 3-s interval, it was then rewarded with water for 3 s. On no-go trials, if a mouse failed to withhold licking within the 3-s interval, it was given a penalty of a gentle air puff to the eyelid. Mice were trained for two sessions per day, with 80 randomized trials (20 min) per session. For analysis, a 'go' response was defined as four or more licks in the second before reward or penalty. For photostimulation experiments, mice were first trained until they could effectively discriminate the tastants with ~90% accuracy (over 1–2 weeks). Then, on the 'probe' sessions, tastants and/or cortical photostimulation were presented during the sample period. Neither reward nor punishment was delivered for novel tastants or light stimulation. Before testing, animals with correctly placed cannulae were provisionally identified by ChR2 expression followed by one or two sessions of lick preference pre-tests. All animals analysed had histologically confirmed placement of cannulae and expression of ChR2 in the appropriate cortical field.

Pharmacological inhibition. Mice were trained to discriminate sweet from bitter in the go/no-go task with at least 90% accuracy. On the day of the experiment, mice were first tested with four taste stimuli (pre-test), including the original training tastants (2 mM AceK and 0.1 mM quinine) and a novel sweet and bitter tastant (50 mM sucrose and 2 μM cycloheximide). After the test, 0.3 μl of the glutamate receptor antagonist NBQX (5 mg ml^{-1} in 0.9% NaCl, Tocris Bioscience) was bilaterally infused into the chosen insular cortical fields over a period of 3 min. NBQX was delivered via an internal infusion needle inserted into the same guide cannulae used for light stimulation and connected to a 1-μl Hamilton syringe (PlasticsOne). Saline (0.9% NaCl) was used as control. After NBQX or saline infusion, animals were placed in their home cages to rest for 1.5 h. Mice were then re-tested with the same four taste stimuli on the go/no-go task (NBQX-test) and then at 8–24 h after rest (recovery-test). During tests, a water reward was given for correctly identifying the go cue, but no air puff was delivered for incorrectly identifying the no-go cue (to avoid possible re-learning). No reward or punishment was applied for the novel sweet and bitter tastants. A performance ratio was calculated for each taste quality: ratio = r_1/r_2, where r_1 is the percentage of correct responses during the NBQX-test or recovery-test, and r_2 is the percentage of correct responses during the pre-test. The percentage of correct responses for each taste quality was the average of go (%) for go taste stimuli (for example, quinine and cycloheximide), or the difference between (100 − go (%)) for no-go stimuli (for example, AceK and sucrose). All animals analysed had anatomically confirmed placement of cannulae in the appropriate cortical field. We note that we made several unsuccessful attempts to optogenetically silence the sweet and bitter cortical fields; this may have been due in part to the requirement for expression in most, if not all, relevant neurons.

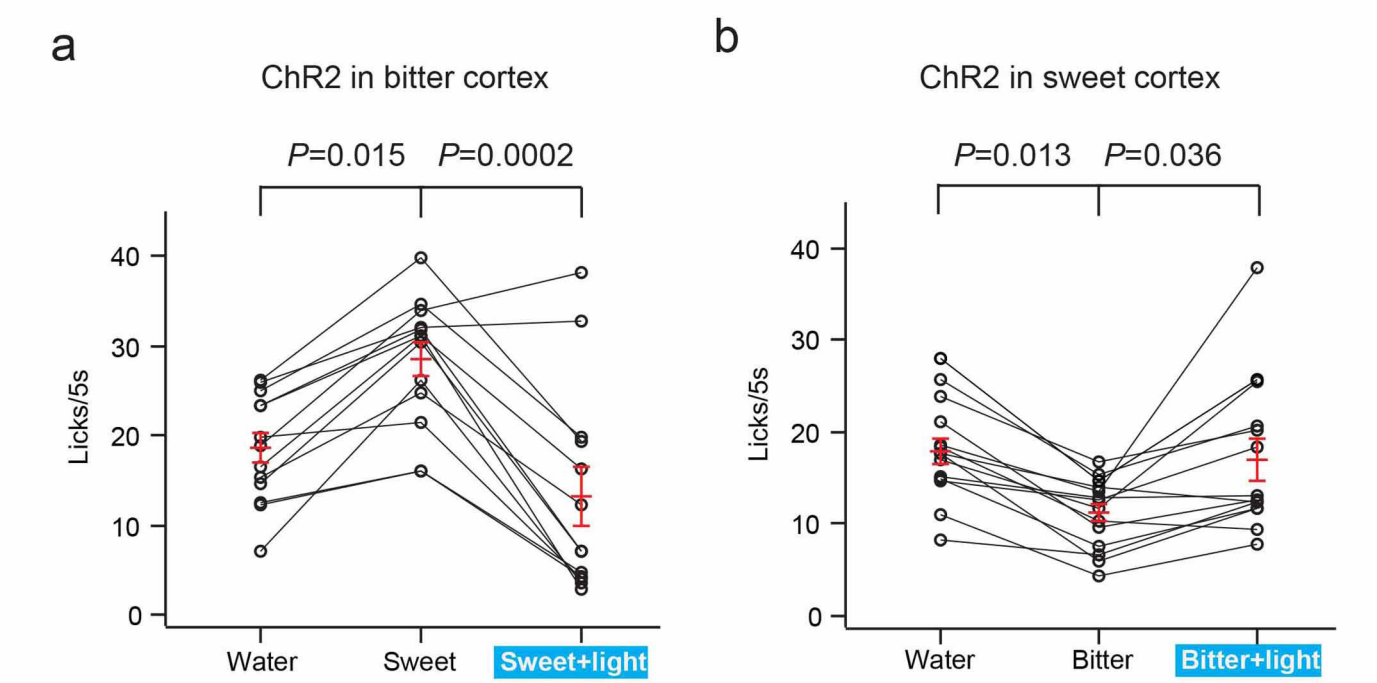
28. Tokita, K., Armstrong, W. E., St John, S. J. & Boughter, J. D. Jr. Activation of lateral hypothalamus-projecting parabrachial neurons by intraorally delivered gustatory stimuli. *Front. Neural Circuits* **8,** 86 (2014).



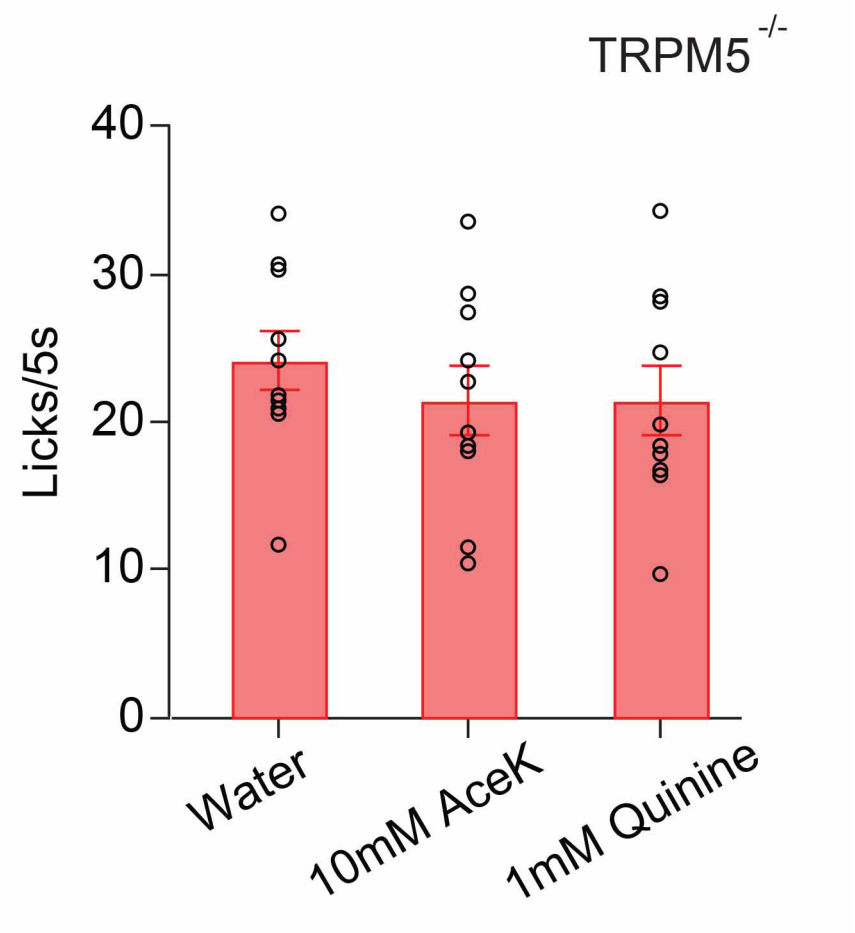
Extended Data Figure 1 | Expression of ChR2 in taste cortex. a, Samples of injection sites in the bitter and sweet cortical fields; shown are coronal sections (Fig. 1a shows a whole mount brain). ChR2–YFP expression (green), nuclei (blue; TO-PRO-3); numbers indicate position relative to bregma, and the dotted area highlight the location of the taste cortical fields (see **c**). **b**, Activation of insular neurons in sweet cortex triggers robust c-Fos expression; ChR2–YFP (green), c-Fos (red) after 10 min of *in vivo* photostimulation at 20 Hz, 20-ms pulses (5 s laser on, 5 s laser off, 5 mW). Dashed lines indicate the location of the stimulating cannulae/fibre. **c**, c-Fos (red) expression in bitter cortex (bregma 0, −0.2) after bitter tastant stimulation (10 mM quinine; see Methods for details). Note the absence of c-Fos expression in the middle (bregma +0.7) and sweet insular cortex (bregma +1.5). Importantly, specific labelling is abolished in taste blind animals (TRPM5 knockouts; middle row). The bottom row shows a diagram of the corresponding brain areas, adapted from the Allen Brain Atlas. Scale bars: 1 mm (**a**), 500 μm (**b**), 300 μm (**c**). PIR, piriform cortex; IC, insular cortex.



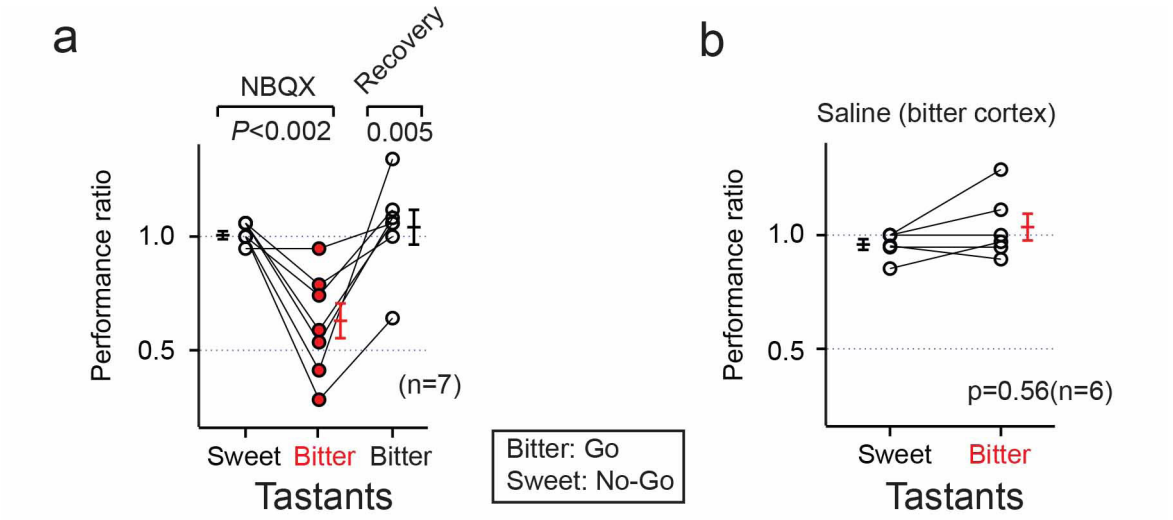
Extended Data Figure 2 | Acquisition of Place preference. a, The development of 'place preference' as a function of session number (each session was 30 min of training and 5 min of 'after-training' testing in the absence of light stimulation; $n = 13$ for sweet cortex, $n = 15$ for bitter cortex; see text and Methods for details). The average of sessions 6–8 was used in Fig. 1. Values are mean ± s.e.m. **b**, Representative mouse track and quantitation of preference index in control GFP-expressing mice; note no difference in preference between chambers ($n = 14$; Mann–Whitney U-test, $P = 0.74$). Values are mean ± s.e.m.



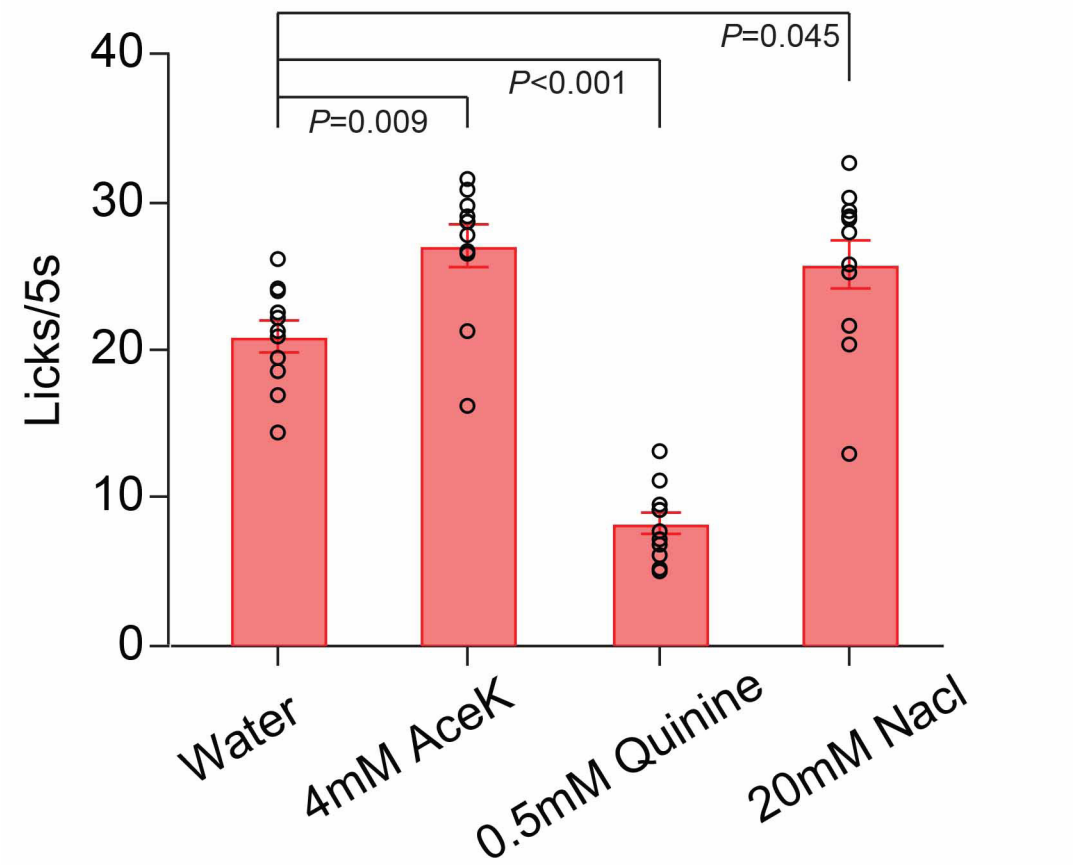
Extended Data Figure 3 | Photostimulation of insular cortical fields overcomes natural taste valence. a, Quantitation of licking responses in mice expressing ChR2 in the bitter cortical fields ($n = 13$, analysis of variance (ANOVA) test, Tukey's honest significant difference post hoc test). Photostimulation of the bitter cortical fields significantly suppress the natural attraction of the sweet tastant (4 mM AceK). **b**, Quantitation of licking responses in mice expressing ChR2 in the sweet cortical fields ($n = 14$, ANOVA test, Tukey's honest significant difference post hoc test). Photostimulation of the sweet cortical fields significantly overcomes the natural aversion of the bitter tastant (1 mM quinine). In both experiments, mice were water-restrained (but exhibited an average of not more than 30 licks per 5-s water trial) such that they were motivated to drink the bitter while showing attraction to sweet. Values are mean ± s.e.m.



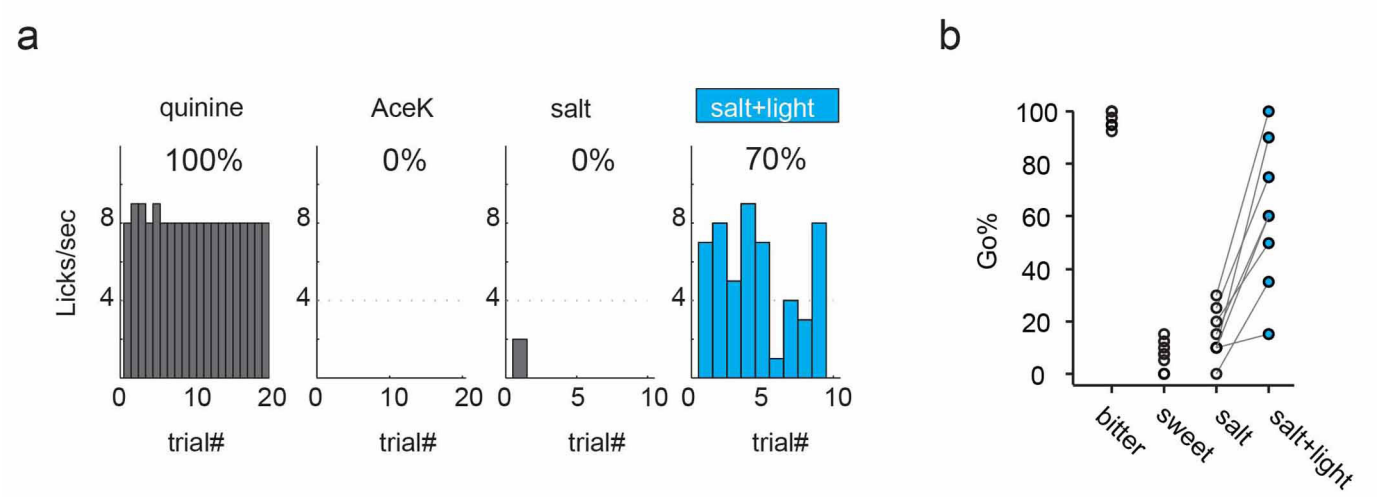
Extended Data Figure 4 | TRPM5 knockout mice do not taste sweet and bitter. Taste preference was tested in the head-restrained assay for wild type and TRPM5 homozygous mutants. Tastants were randomly delivered for a 5-s window (ten trials each). No significant difference was observed between water and sweet/bitter tastants in TRPM5 knockouts (ANOVA test, $P = 0.62$, $n = 10$; see ref. 10 for more details); circles indicate individual animals; bar graphs show mean ± s.e.m.



Extended Data Figure 5 | Inactivation of the bitter cortical fields in animals trained to go to bitter and no-go to sweet. a, Quantitation of performance ratios before and after bilateral silencing of the bitter cortical fields (NBQX, 5 mg ml^{-1}; $n = 7$) in animals trained to go to bitter and no-go to sweet. Note the impact in bitter taste discrimination, but no significant effect in sweet taste (Mann–Whitney *U*-test, $P < 0.002$). After washout of the drug, the animal's ability to recognize bitter is restored (Mann–Whitney *U*-test, $P < 0.005$). **b**, Quantitation of performance ratios with saline (0.9%) control in the bitter cortical fields ($n = 6$, Mann–Whitney *U*-test, $P = 0.56$). In both experiments, mice were trained with quinine and AceK, and tested with two pairs of sweet/bitter tastants (0.1 mM quinine and 2 mM AceK, 2 μM cycloheximide and 50 mM sucrose; see Methods for details).



Extended Data Figure 6 | Sweet and low salt are appetitive tastants. Taste preference was tested during a 10-min window using the head-restrained assay (see Methods for details). Four tastants were randomly delivered to animals for 5 s each (ten trials per tastant). Note that animals show significant attraction to sweet (AceK) and low salt (NaCl), but strong aversion to bitter ($n = 11$, ANOVA test, Tukey's honest significant difference post hoc test); circles indicate individual animals; bar graphs show mean ± s.e.m. These conditions were used in the experiments described in Fig. 5 and Extended Data Fig. 7.



Extended Data Figure 7 | Cross-generalization between orally supplied taste stimuli and photostimulation of the bitter cortex. a, Representative histograms illustrating cross-generalization between taste stimulation and photostimulation of the bitter cortical field. The mouse was trained to go to bitter (0.5 mM quinine) and no-go to sweet (4 mM AceK) and low salt (20 mM NaCl). **b**, Quantitation of the responses from individual animals to quinine, AceK, salt and salt + light ($n = 8$, Mann–Whitney U-test, $P < 0.002$). See also Fig. 4.

LETTER

doi:10.1038/nature16148

Drosophila Ionotropic Receptor 25a mediates circadian clock resetting by temperature

Chenghao Chen[1]*, Edgar Buhl[2]*, Min Xu[1], Vincent Croset[3], Johanna S. Rees[4], Kathryn S. Lilley[4], Richard Benton[3], James J. L. Hodge[2] & Ralf Stanewsky[1]

Circadian clocks are endogenous timers adjusting behaviour and physiology with the solar day[1]. Synchronized circadian clocks improve fitness[2] and are crucial for our physical and mental well-being[3]. Visual and non-visual photoreceptors are responsible for synchronizing circadian clocks to light[4,5], but clock-resetting is also achieved by alternating day and night temperatures with only 2–4 °C difference[6–8]. This temperature sensitivity is remarkable considering that the circadian clock period (~24 h) is largely independent of surrounding ambient temperatures[1,8]. Here we show that *Drosophila* Ionotropic Receptor 25a (IR25a) is required for behavioural synchronization to low-amplitude temperature cycles. This channel is expressed in sensory neurons of internal stretch receptors previously implicated in temperature synchronization of the circadian clock[9]. IR25a is required for temperature-synchronized clock protein oscillations in subsets of central clock neurons. Extracellular leg nerve recordings reveal temperature- and IR25a-dependent sensory responses, and IR25a misexpression confers temperature-dependent firing of heterologous neurons. We propose that IR25a is part of an input pathway to the circadian clock that detects small temperature differences. This pathway operates in the absence of known 'hot' and 'cold' sensors in the *Drosophila* antenna[10,11], revealing the existence of novel periphery-to-brain temperature signalling channels.

In *Drosophila*, daily activity rhythms are controlled by a network of ~150 clock neurons expressing the clock genes *period* (*per*) and *timeless* (*tim*). These encode repressor proteins that negatively feedback on their own promoters resulting in 24 h oscillations of clock molecules. Temperature cycles (TC) synchronize molecular clocks present in peripheral appendages in a tissue-autonomous manner[9,12], whereas synchronization of clock neurons in the brain mainly depends on peripheral temperature receptors located in the chordotonal organs (ChO) and the ChO-expressed gene *nocte*[9,12,13].

To discover novel factors involved in temperature entrainment, we identified NOCTE-interacting proteins by co-immunoprecipitation and mass-spectrometry (Extended Data Table 1)[14]. We focused on IR25a, a member of a divergent subfamily of ionotropic glutamate receptors and verified the interaction by co-immunoprecipitation after overexpressing IR25a and NOCTE in all clock cells using *tim-gal4* (Extended Data Fig. 1a). IR25a is expressed in different populations of sensory neurons, including those in the antenna and labellum[15–17]. In the olfactory system IR25a acts as a co-receptor with different odour-sensing IRs[15].

To investigate if *IR25a* is co-expressed with *nocte* in ChO, we analysed *IR25a* expression in femur and antennal ChO using an *IR25a-gal4* line[15] (Extended Data Fig. 2a). *IR25a-gal4*-driven *mCD8–GFP* labelled subsets of ChO neurons in the femur, overlapping substantially with *nompC–QF* driven *QUAS–Tomato* signals (using the *QF* binary transcriptional activation system) (Fig. 1a–c). *nompC–QF* is expressed in larval ChO[18] and in the adult femur ChO (Fig. 1d, e). Comparison of *IR25a*-driven *mCD8–GFP* and nuclear *DsRed* signals with those of other ChO neuron drivers (*F-gal4* and *nocte-gal4* (ref. 9)) suggests that *IR25a* is expressed in a subset of femur ChO neurons and Johnston's Organ (JO) neurons (Fig. 1c and Extended Data Fig. 1b–g). To determine if *IR25a-gal4* ChO signals reflect endogenous IR25a expression, we confirmed the presence of *IR25a* mRNA in the femur and leg (Extended Data Fig. 2b, e) and the co-localization of anti-IR25a immunofluorescence signals in femur ChO neurons (Fig. 1f, g). IR25a was detected in ChO neuron cell bodies and ciliated dendrites, as was an mCherry–IR25a fusion protein expressed in these cells (Fig. 1h).

As *nocte*[1] mutants do not synchronize to 12 h:12 h 16 °C:25 °C temperature cycles in constant light (LL)[9,12] (Extended Data Fig. 3a), we analysed $IR25a^{-/-}$ mutants[16] under these conditions. Unlike *nocte*[1], the $IR25a^{-/-}$ flies synchronized well to this regime and we obtained similar results at warmer temperature cycles (Extended Data Fig. 3a). To test whether *IR25a* is specifically required for synchronization to small temperature intervals[7,13], we subjected $IR25a^{-/-}$ flies to various temperature cycles with an amplitude of only 2 °C. Surprisingly, and in contrast to wild-type, $IR25a^{-/-}$ mutants did not synchronize to any of the shallow temperature cycles in LL or constant darkness (DD) (Fig. 2a–e and Extended Data Figs 3b and 4c). In LL, wild-type and *IR25a* rescue flies showed a clear activity peak in the second part of the warm period before and after the 6 h shift of the temperature cycle. By contrast, $IR25a^{-/-}$ mutants were constantly active throughout the temperature cycle, apart from a short period of reduced activity at the beginning of the warm phase of TC1 (Fig. 2a and Extended Data Fig. 3b). In DD, control flies slowly advanced (or delayed) their evening activity peak during phase-advanced (or delayed) temperature cycles (Fig. 2b and Extended Data Fig. 4c). The phase of this activity peak was maintained in the subsequent free-running conditions (DD, constant 25 °C) indicating stable re-entrainment of the circadian clock (Fig. 2b and Extended Data Fig. 4). By contrast, *IR25a* mutants did not shift their evening peak during the temperature cycle, keeping their original phase throughout the experiment (Fig. 2b and Extended Data Fig. 4c).

To quantify entrainment in LL, we determined the 'entrainment index' (EI)[13], whereas for most DD experiments we calculated the phase difference of the main activity peak upon release into constant conditions between *IR25a* mutants and controls. In all 2 °C amplitude temperature cycles tested the entrainment index of $IR25a^{-/-}$ flies was significantly lower and phase calculation indicated no phase shift or a significantly reduced phase shift compared to controls (Fig. 2c–e). The same non-synchronization phenotype was observed in $IR25a^{-/}$ *Df(IR25a)* flies, and temperature synchronization was fully restored in $IR25a^{-/-}$ rescue flies (Fig. 2a–d and Extended Data Fig. 3b). $IR25a^{-/-}$ mutants synchronize to light and have normal free-running and

[1]Department of Cell and Developmental Biology, University College London, 21 University Street, London WC1E 6DE, UK. [2]School of Physiology and Pharmacology, University of Bristol, University Walk, Bristol BS8 1TD, UK. [3]Center for Integrative Genomics, Faculty of Biology and Medicine, University of Lausanne, CH-1015 Lausanne, Switzerland. [4]Cambridge Centre for Proteomics, Department of Biochemistry and Cambridge Systems Biology Centre, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, UK.
*These authors contributed equally to this work.

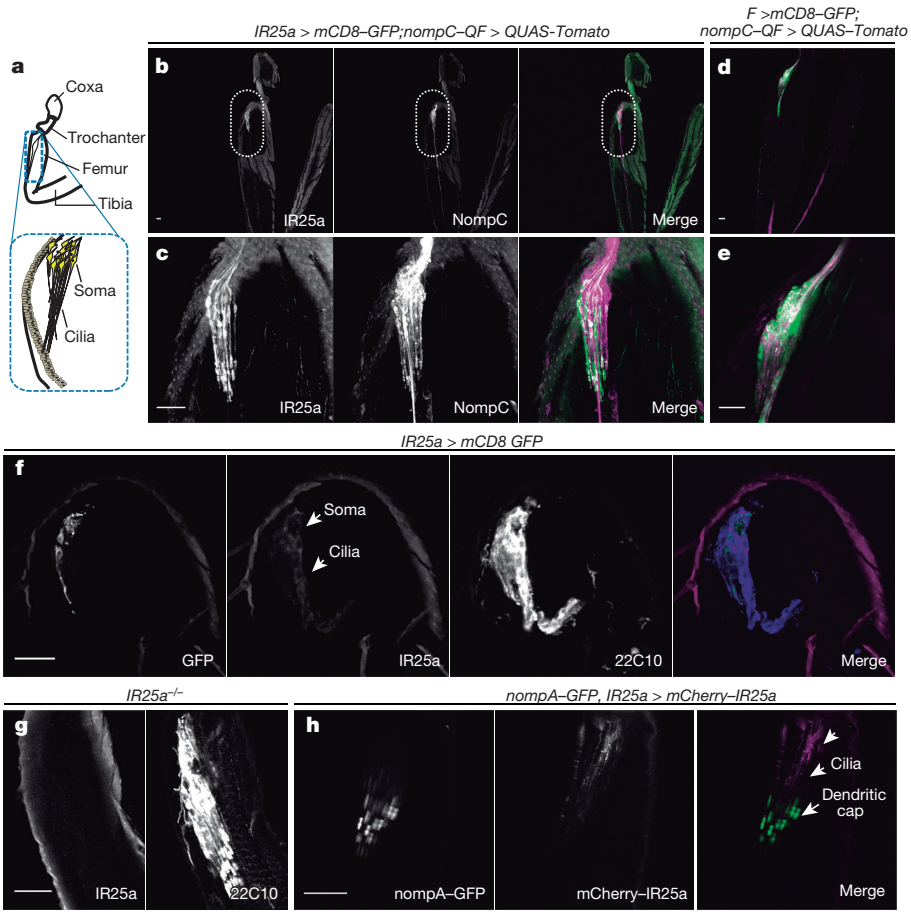


Figure 1 | IR25a is expressed in ChO neurons. **a**, Overview of the femur ChO adapted from[13]. **b**, **d**, Double labelling of the femur ChO by *IR25a-gal4* (**b**) and *F-gal4* (**d**) driven mCD8-GFP and *nompC-QF* driven *QUAS-Tomato*. **c**, **e**, Higher magnification of circled ChO areas in **b** and **d**, respectively. **f**, IR25a immunolabelling of femoral ChO cryosections of *IR25a-gal4/UAS-mCD8-GFP* flies. From left to right, GFP, anti-IR25a, 22C10, and merged images are shown. **g**, Anti-IR25a and 22C10 labelling of femur ChO sections of $IR25a^{-/-}$ flies. **h**, Subcellular distribution of an mCherry–IR25a fusion protein co-labelled with the dendritic cap marker nompA–GFP in the femur ChO. Scale bar, 20 μm.

temperature compensated periods (Fig. 2b, Extended Data Fig. 4d and Extended Data Table 2). These results suggest that IR25a enables the circadian clock to sense subtle temperature changes across the entire physiological range, rather than mediating synchronization to a specific range. Increasing the temperature cycle amplitude to 4 °C consistently restored temperature entrainment in $IR25a^{-/-}$ flies (Extended Data Fig. 4a, b).

Temperature receptors located in fly antennae and arista are not required for temperature-synchronized behaviour[9,11,19]. As expected, we found that antennal *IR25a* function (Extended Data Figs 1c and 2a)[16] is not required for temperature entrainment (Extended Data Fig. 5). To reveal the importance of IR25a expression in ChO neurons, we performed tissue-specific *IR25a* RNA interference (RNAi) using validated transgenes (Extended Data Figs 2d and 6a). *IR25a* RNAi in all or subsets of ChO neurons (Fig. 1 and Extended Data Fig. 1) resulted in a lack of entrainment (Extended Data Figs 2e and 6b, c). By contrast, *IR25a* RNAi in multidendritic, TRPA1-expressing or clock neurons did not impair temperature entrainment (Extended Data Fig. 6c). These findings are consistent with the absence of *IR25a* expression in clock neurons and the brain (Extended Data Fig. 2e–g) and show that *IR25a* functions in ChO neurons for temperature entrainment to 25 °C:27 °C temperature cycles in LL.

To identify the neural substrates underlying the lack of behavioural synchronization, we quantified clock protein levels in wild-type, $IR25a^{-/-}$, and $IR25a^{-/-}$ rescue flies exposed to a shallow temperature cycle in LL. Although TIM expression was robustly rhythmic and synchronized in all clock neuronal groups in controls, TIM was barely detectable in the Dorsal Neuron 1 (DN1) and DN2 of $IR25a^{-/-}$ flies (Fig. 3a and Extended Data Fig. 7a, b). Moreover, in the small and large ventral lateral neurons (s-LNv and l-LNv), TIM expression exhibited an additional peak during the warm phase (Fig. 3a and Extended Data Fig. 7a, b). In the DN3, TIM declined earlier compared to controls and there was no effect on the dorsal lateral neurons (LNd). In temperature cycles and DD, TIM levels in DN1 were also blunted but oscillations in the DN2 and DN3 were similar to controls. In contrast to LL, TIM did not oscillate in any of the LN groups and was at constantly low levels (Fig. 3b), consistent with the behavioural results obtained under these conditions (Fig. 2b, d). The alterations of TIM expression are temperature specific, as we observed normal oscillations in LD cycles at 25 °C (Extended Data Fig. 7c). An increase of the temperature cycle amplitude to 4 °C also restored normal TIM expression in $IR25a^{-/-}$ flies, in agreement with the behavioural rescue (Extended Data Figs 4a, b and 7d). In summary, in low-amplitude temperature cycles, IR25a is required for normally synchronized TIM oscillations in DN1-3 and LNv in LL and in DN1 and LN clock neurons in DD.

We tested if the clock neurons affected by the lack of IR25a are indeed involved in regulating behavioural synchronization to shallow temperature cycles by blocking synaptic transmission using tetanus-toxin (TNT). Indeed, TNT-expression in DN1 and DN2 blocked synchronization in LL, whereas in DD only DN1 blockage interfered with temperature entrainment (Fig. 3c, d)[20]. Consistent with the differential effect on TIM oscillations in LL and DD (Fig. 3a, b) these results strongly suggest that IR25a is required for the synchronized output of the DN1 (LL and DD) and DN2 (LL) to control temperature-entrained behaviour.

Next, we asked if ChO might directly sense temperature in an IR25a-dependent manner. We recorded leg nerve activity in restrained preparations and identified ChO units in the compound signal (Fig. 4a). In both wild-type and $IR25a^{-/-}$ flies, spontaneous leg movement changed as a function of temperature along with motor and sensory activity. Additionally, presumed ChO activity of wild-type flies also increased during periods without movement (Fig. 4b, third insert). This temperature-induced but movement-independent, ChO activity was absent in $IR25a^{-/-}$ flies, showing that temperature

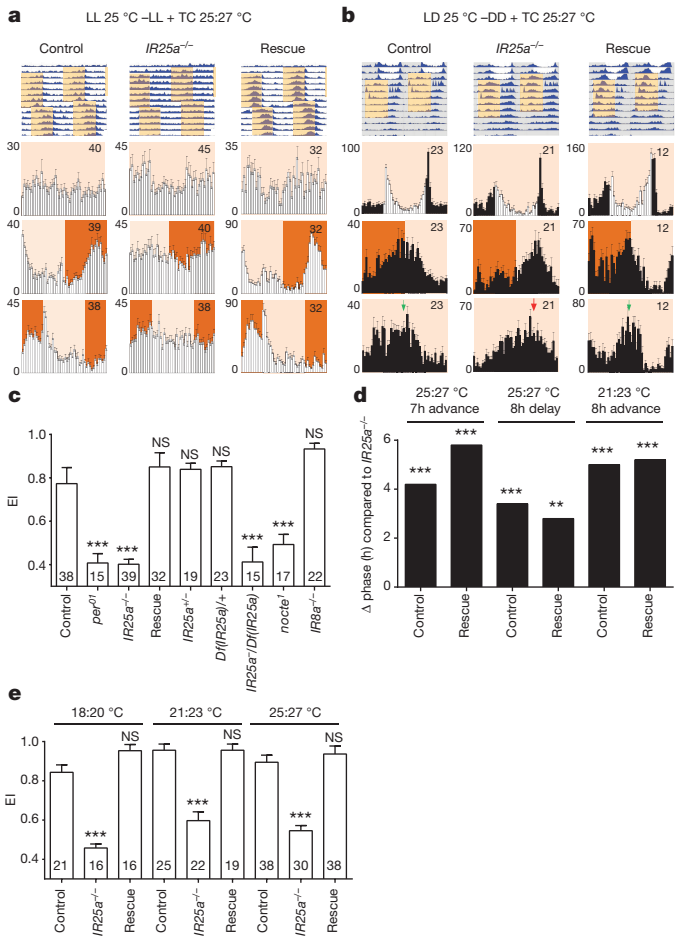


Figure 2 | *IR25a* is required for temperature synchronization to low-amplitude temperature cycles. a, Upper part shows double plotted average actograms depicting the daily activity levels and environmental conditions during the entire experiment. White areas, LL and 25 °C; orange areas, LL and 27 °C. Histograms show daily average activity levels during the initial LL treatment and the last 3 days of each temperature cycle. Light orange, 25 °C; dark orange, 27 °C; white bars, activity levels in LL. Error bars indicate s.e.m.; numbers (*n*) in the upper-right corner; *x* axis, Zeitgeber time (h) and *y* axis total activity (beam crossings per 30 min). **b**, As in **a** but flies were initially kept in LD 25 °C, before being exposed to a 7 h phase advanced temperature cycle in DD (dark histogram bars) and free-running conditions (DD and 25 °C). Actogram shading as in **a** but grey areas indicate darkness. Green and red arrows indicate the position (phase) of the main activity peak during the final free run for control and mutant flies, respectively. **c**, **e**, Entrainment index values (mean ± s.e.m.) during 25 °C:27 °C temperature cycles in LL (delay as in **a**) (**c**), and as indicated in **e** (all delay, except 25 °C:27 °C, advance) (see Extended Data Fig. 3b for actograms and daily average plots). In **c**, per^{01} and $nocte^{1}$ flies were used as negative controls. $^{***}P < 0.001$, $^{**}P < 0.01$, NS, not significant, one way ANOVA followed by Bonferroni correction. **d**, Phase difference during DD and constant temperature after temperature cycles between $IR25a^{-/-}$ ($n = 12/11/12$ for 7 h advance/8 h delay/8 h advance temperature cycles, respectively) and *y w* control ($n = 16/10/14$, respectively) and $IR25a^{-/-}$ rescue flies ($n = 16/18/12$). $^{****}P < 0.0001$, $^{***}P < 0.001$, $^{**}P < 0.01$; *F*-statistic (Watson–Williams–Stevens test).

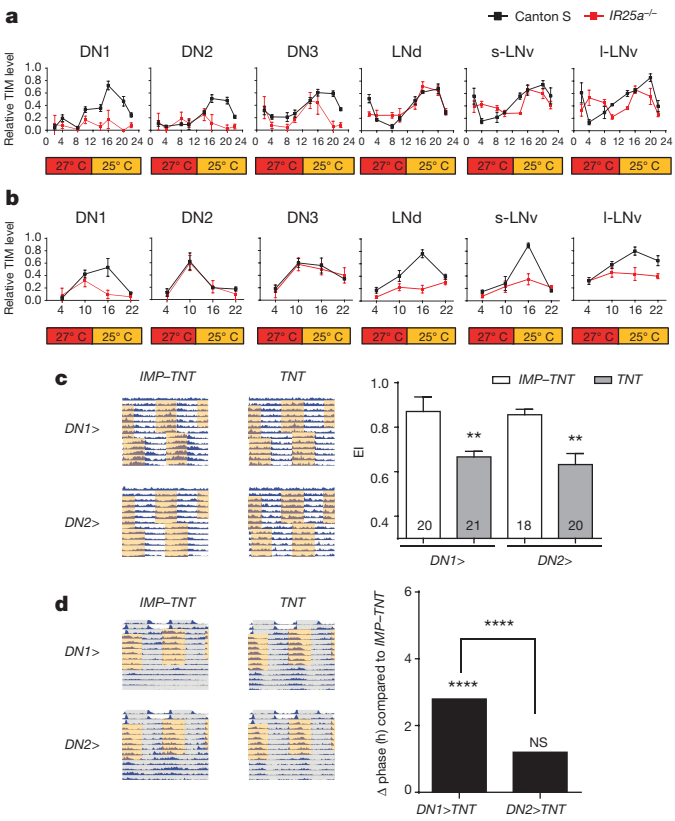


Figure 3 | IR25a is required for clock protein oscillations in central clock neurons. a, **b**, TIM levels in clock neurons during LL (**a**) and DD (**b**) 25 °C:27 °C temperature cycles at the indicated time points Zeitgeber time (ZT). At least 8 brain hemispheres per time point were analysed for each genotype. Error bars indicate s.e.m. **c**, Progeny of *UAS-IMP-TNT* and *UAS-TNT* females crossed to *Clk4.1M-gal4* (DN1>, upper panel) or *Clk9M-gal4;Pdf-gal80* (DN2>, lower panel) males, were exposed to two 6 h-delayed temperature cycles (12 h at 25 °C:12 h at 27 °C in LL). Left, actograms, shading as in Fig. 2a. Right, entrainment index calculations (mean ± s.e.m.), numbers in bars indicate *n*. $^{**}P < 0.01$; One way ANOVA followed by Bonferroni correction. **d**, Same genotypes as in **c** were exposed to an 8 h delayed 12 h 25 °C:12 h 27 °C temperature cycles in DD. Left: actograms plotted as in Fig. 2b, Right: phase difference of activity peaks during final constant conditions between controls (DN1/DN2 > *UAS-IMP-TNT*, $n = 9/12$, respectively) and the indicated genotypes (DN1/DN2 > TNTE, $n = 16/10$). $^{****}P < 0.0001$, NS, not significant, *F*-statistic (Watson–Williams–Stevens test).

is sensed in the legs in an IR25a-dependent manner (Fig. 4c). To test if IR25a contributes directly to temperature-sensing, we ectopically expressed this channel in the physiologically well-characterized, IR25a-negative, l-LNv (Extended Data Fig. 2f). As a positive control, we also expressed the temperature-sensitive *Drosophila* TRPA1 channel[21] in the l-LNv. Isolated brains were exposed to a temperature ramp, and spike frequency of individual l-LNv was recorded. Control l-LNv did not show a significant temperature-dependent change in neural activity (Fig. 4d). As expected, the firing rate of TRPA1 expressing neurons drastically increased linearly with temperature, as did other cellular parameters (Extended Data Fig. 8). *IR25a* expression resulted in a linear and reversible temperature-dependent increase in action potential firing frequency (Fig. 4e, i), whereas other cellular parameters showed no difference (Fig. 4f–h). Increasing the temperature by only 2–3 °C also lead to a reversible increase in firing frequency of 1.03 ± 0.20 Hz in *IR25a* expressing l-LNv (Fig. 4j). By contrast, expression of the related, but olfactory-specific co-receptor *IR8a* (which is not required for temperature entrainment, Fig. 2c) did not confer temperature-sensitivity (Extended Data Fig. 8). These observations suggest that IR25a is at least part of a thermosensory receptor required for temperature entrainment.

Our data indicating that IR25a contributes to temperature sensing within ChO extend the roles of IR's beyond chemoreception, reminiscent of the requirement for the 'gustatory receptor' Gr28b in warmth-avoidance[22]. Although we show that IR25a-expressing leg neurons are capable of sensing temperature and mediating temperature entrainment, it is possible that this receptor has a similar role elsewhere in the peripheral nervous system. IR25a responds to small temperature changes and we propose that the fly continuously integrates

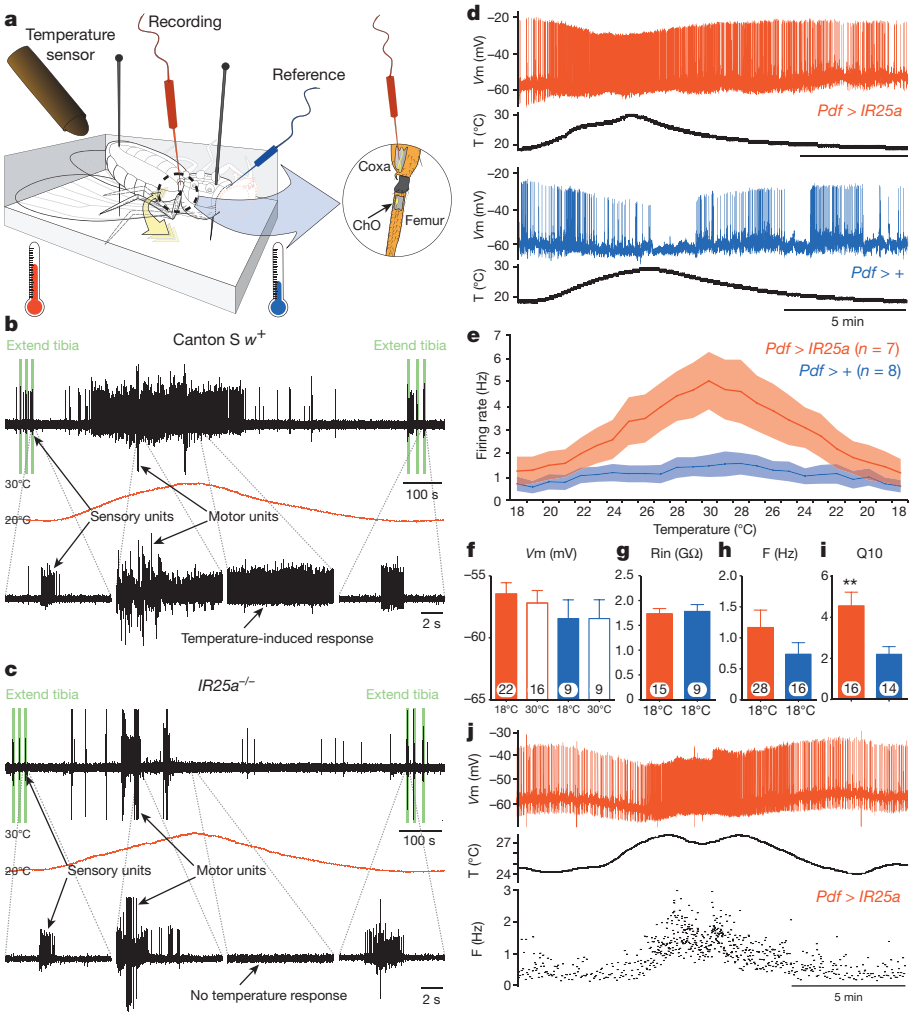


Figure 4 | IR25a is required for temperature-induced leg nerve responses and confers temperature sensitivity to l-LNv. a, Schematic of the setup. **b**, Recording of a control fly leg nerve including motor and sensory axons. The first extended insert shows a discharge of presumed ChO sensory units in response to manual extension of the tibia (green bars). Heating the preparation from 20 °C to 30 °C (middle, red trace) lead to spontaneous leg movement with concurrent motor and sensory activity (second insert) but also to increased sensory firing in the absence of leg or motor activity (third insert), which was reversible with intact tibia extension response (fourth insert) ($n = 9$). **c**, *IR25a*$^{-/-}$ shows similar responses to tibia extension and temperature-dependent leg movement, but no sensory activity in response to elevated temperature ($n = 6$). **d**, Whole-cell current clamp recordings of l-LNv control and *Pdf > IR25a* brains exposed to the indicated temperature ramp. **e**, Quantification of the temperature response from multiple recordings (mean, s.e.m.). **f**, *V*m, membrane potential. **g**, Rin, input resistance. **h**, F, spontaneous firing rate at 18 °C. **i**, Temperature coefficient Q10 > 4, $^{**}P < 0.01$; *t*-test. Error bars indicate s.e.m., numbers (*n*) written on the bars. **j**, *IR25a* expressing l-LNv also report small (2–3 °C) temperature changes leading to a ~1 Hz alteration of the instantaneous firing rate ($n = 5$).

temperature signals received from multiple ChO across the whole body for synchronization of the clock. This potential reliance on weakly responding temperature receptors might explain why the *Drosophila* circadian clock is insensitive to brief temperature pulses[23], which could help maintain synchronized clock function in natural conditions of rapid and large temperature fluctuations.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 14 April; accepted 26 October 2015.
Published online 18 November 2015.

1. Dunlap, J. C., Loros, J. J. & DeCoursey, P. J. *Chronobiology: Biological Timekeeping* (Sinauer Associates, 2004).
2. Ouyang, Y., Andersson, C. R., Kondo, T., Golden, S. S. & Johnson, C. H. Resonating circadian clocks enhance fitness in cyanobacteria. *Proc. Natl Acad. Sci. USA* **95,** 8660–8664 (1998).
3. Bechtold, D. A., Gibbs, J. E. & Loudon, A. S. Circadian dysfunction in disease. *Trends Pharmacol. Sci.* **31,** 191–198 (2010).
4. Helfrich-Förster, C., Winter, C., Hofbauer, A., Hall, J. C. & Stanewsky, R. The circadian clock of fruit flies is blind after elimination of all known photoreceptors. *Neuron* **30,** 249–261 (2001).
5. Hughes, S., Jagannath, A., Hankins, M. W., Foster, R. G. & Peirson, S. N. Photic regulation of clock systems. *Methods Enzymol.* **552,** 125–143 (2015).
6. Brown, S. A., Zumbrunn, G., Fleury-Olela, F., Preitner, N. & Schibler, U. Rhythms of mammalian body temperature can sustain peripheral circadian clocks. *Curr. Biol.* **12,** 1574–1583 (2002).
7. Wheeler, D. A., Hamblen-Coyle, M. J., Dushay, M. S. & Hall, J. C. Behavior in light-dark cycles of *Drosophila* mutants that are arrhythmic, blind, or both. *J. Biol. Rhythms* **8,** 67–94 (1993).
8. Maguire, S. E. & Sehgal, A. Heating and cooling the *Drosophila melanogaster* clock. *Curr. Opin. Insect Sci.* **7,** 71–75 (2015).
9. Sehadova, H. *et al.* Temperature entrainment of *Drosophila*'s circadian clock involves the gene *nocte* and signaling from peripheral sensory tissues to the brain. *Neuron* **64,** 251–266 (2009).
10. Florence, T. J. & Reiser, M. B. Neuroscience: hot on the trail of temperature processing. *Nature* **519,** 296–297 (2015).
11. Gallio, M., Ofstad, T. A., Macpherson, L. J., Wang, J. W. & Zuker, C. S. The coding of temperature in the *Drosophila* brain. *Cell* **144,** 614–624 (2011).
12. Glaser, F. T. & Stanewsky, R. Temperature synchronization of the *Drosophila* circadian clock. *Curr. Biol.* **15,** 1352–1363 (2005).
13. Wolfgang, W., Simoni, A., Gentile, C. & Stanewsky, R. The Pyrexia transient receptor potential channel mediates circadian clock synchronization to low temperature cycles in *Drosophila melanogaster*. *Proc. R. Soc. Lond. B* **280,** 20130959 (2013).
14. Rees, J. S. *et al. In vivo* analysis of proteomes and interactomes using parallel affinity capture (iPAC) coupled to mass spectrometry. *Mol. Cell Proteomics* **10,** M110.002386 (2011).
15. Abuin, L. *et al.* Functional architecture of olfactory ionotropic glutamate receptors. *Neuron* **69,** 44–60 (2011).
16. Benton, R., Vannice, K. S., Gomez-Diaz, C. & Vosshall, L. B. Variant ionotropic glutamate receptors as chemosensory receptors in *Drosophila*. *Cell* **136,** 149–162 (2009).
17. Rytz, R., Croset, V. & Benton, R. Ionotropic receptors (IRs): chemosensory ionotropic glutamate receptors in *Drosophila* and beyond. *Insect Biochem. Mol. Biol.* **43,** 888–897 (2013).
18. Petersen, L. K. & Stowers, R. S. A Gateway MultiSite recombination cloning toolkit. *PLoS ONE* **6,** e24531 (2011).
19. Sayeed, O. & Benzer, S. Behavioral genetics of thermosensation and hygrosensation in *Drosophila*. *Proc. Natl Acad. Sci. USA* **93,** 6079–6084 (1996).
20. Kaneko, H. *et al.* Circadian rhythm of temperature preference and its neural control in *Drosophila*. *Curr. Biol.* **22,** 1851–1857 (2012).
21. Hamada, F. N. *et al.* An internal thermal sensor controlling temperature preference in *Drosophila*. *Nature* **454,** 217–220 (2008).
22. Ni, L. *et al.* A gustatory receptor paralogue controls rapid warmth avoidance in *Drosophila*. *Nature* **500,** 580–584 (2013).
23. Busza, A., Murad, A. & Emery, P. Interactions between circadian neurons control temperature synchronization of *Drosophila* behavior. *J. Neurosci.* **27,** 10722–10733 (2007).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank P. Emery, J. Albert, J. Jepson, P. Garrity, and A. Samuel for discussions and sharing of unpublished results, J. Giebultowicz for anti-TIM antibodies, J. Albert and J. Jepson for fly stocks, C. Tardieu and R. Kavlie for help with qPCR, D. Carr for assistance with the temperature recording setup, and M. Ogueta-Gutierrez for help with figure preparations. The drawing for Fig. 4a was generated by Polygonal Tree (http://polygonaltree.co.uk/). This work was supported by BBSRC grants BB/H001204 to R.S., BB/J0-18589/-17221 to R.S. and J.J.L.H., and a CSC PhD fellowship to C.C. V.C. was supported by a Boehringer Ingelheim Foundation Fellowship. Research in R.B.'s laboratory was supported by European Research Council Starting Independent Researcher and Consolidator Grants (205202 and 615094). Mass spectrometry analysis was supported by Wellcome Trust grant 099135/Z/12/Z.

Author Contributions C.C., E.B., R.S., J.J.L.H., R.B. and K.S.L. conceived, designed, and supervised the project. C.C., E.B., M.X., V.C., and J.S.R. performed experiments. C.C., E.B., and J.S.R. analysed data, and R.S. wrote the paper, with feedback from all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.S. (r.stanewsky@ucl.ac.uk).

METHODS

Plasmids and germline transformations. To generate the *psp-flag-strep II-nocte-ha* (*FSNH*) construct, a *flag-strepII-venus-strepII (fsvs)* fragment was amplified from a *PiggyBac/P-element YFP-flag-strep II* construct of[14] using a Phusion High-Fidelity PCR kit (New England Biolabs). This 900 bp fragment was sub-cloned into *psp73* (Promega) to generate *psp73-fsvs* with BglII/XhoI sites. To introduce a *strepII* tag upstream of the NOCTE N terminus, a 0.5 kb fragment was amplified from *psp73-nocte-HA* (containing the entire *nocte* coding region fused to *ha*; Giesecke and Stanewsky, unpublished) by annealing the *strep II* tag directly using PCR. This fragment was religated back into *psp73-nocte-ha* to generate *psp-strep II-nocte-ha.* A 3× Flag tag was introduced 5′ of *psp73-strep II-nocte-ha* by sub-cloning the *strep II-nocte-ha* fragment into *psp-fsvs* with BstBI/XhoI sites replacing *venus*. To generate the *psp-flag-strep II-nocte-strep II* construct (*FSNS*), a Strep II tag was amplified and annealed 3′ of *psp-nocte* (Giesecke and Stanewsky, unpublished) to generate *psp-nocte-strep II*, followed by sub-cloning into *psp-fsnh* using BstBI/XhoI sites. *FSNH* and *FSNS* were sub-cloned into the transformation vector *pUAST* using BglII/XhoI sites, and transgenic flies were generated using classical transposase-mediated germline transformation. To generate *mCherry–IR25a*, the coding sequence of IR25a lacking the endogenous signal sequence (starting from codon 31) was PCR amplified, subcloned into *pUAST-mCherry attB*[15], and integrated into attP2. To generate the *IR25a* genomic rescue construct, the bacterial artificial chromosome (BAC) CH322-32C20[24] was integrated into attP16, and then recombined onto the *IR25a*^{2} mutant chromosome. Restoration of IR25a expression by this BAC (in *IR25a*^{2}*, CH322-32C20/IR25a*^{2}*, CH322-32C20* animals) was verified by immunostaining with anti-IR25a antibodies (data not shown). All constructs generated in this study were confirmed by DNA sequencing.

Fly strains. Flies were kept at 25 °C or 18 °C on common cornmeal-yeast-sucrose food under light:dark cycles and 60–70 % humidity. As controls, wild-type Canton S and *y w* flies (both carrying the *ls-tim* allele) were used. The following flies used in this study were previously described or obtained from the Bloomington Stock Center: *tim-gal4:67* (ref. 25), *Clock856-gal4* (ref. 26), *F-gal4:33-5* (ref. 27), *nocte-gal4* (ref. 9) *IR25a-gal4* (ref. 15), *gmr-gal4* (BL1104), *Pdf-gal4* (ref. 28), *elav-gal4; UAS-dicer* (BL25750), *UAS-dicer* (BL24646), *trpA1-gal4* (BL27593), *nompC-gal4* (ref. 29), *nompC-QF* (BL36346), *ppk-gal4* (BL32078), *UAS-GFP* (ref. 25), *nompA-GFP*[30], *UAS-mCherry* (BL52268), *QUAS-mtdTomato* (BL30037), *UAS-mCD8-GFP;UAS-DsRed*[9], *Pdf-RFP*[31], *UAS-TrpA1* (ref. 21)*UAS-IR25a*[15], *y per*^{01} *w* and *y per*^{L} *w*[32], *IR25a*$^{-/-}$: either homozygous *IR25a*^{2} or *IR25a*^{2}*/IR25a*^{1} flies, both null mutant alleles generated by gene targeting[16], *IR25a*^{2}*, CH322-32C20/IR25a*^{2}*, CH322-32C20* (ref. 15) (outcrossed to Canton S for six generations and here referred to as *IR25a* rescue), *IR8a*^{1}: null allele of *IR8a* and referred to as *IR8a*$^{-/-}$ [15]*, nocte*^{1}: encodes truncated version of the NOCTE protein[9], *UAS-TNT-E* and *UAS-IMP-TNT-V1-B* (inactive)[33]. *Clk4.1M-gal4* and *Clk9M-gal4;Pdf-gal80* flies were used to direct GAL4 expression to subsets of the DN1p and to the DN2, respectively[20,34]. *IR25a-RNAi* lines 15627-R1 and 15627-R2 were obtained from the NIG-Fly Stock Center. *w*^{1118}; *Df(2L)Exel6010/CyO* was used as *IR25a* deficiency (BL7496).

Immunostaining and quantification. GFP and/or RFP signals were analysed as described in[9]. Briefly, antennae and legs were fixed and dissected in 4% paraformaldehyde/PBS solution. Samples were then washed 3 times in 3% PBST at room temperature followed by mounting in Vectashield (Vector Labs) medium and inspected using a Leica TCS SP5 confocal microscope. To visualize endogenous IR25a expression in the ChO of fly antennae and legs, cryosections (16 μm) and immunolabelling were performed as described in[35] with minor modifications: sections were collected on slides and fixed for 10 min in 4 % formaldehyde in PBS. After washing for 2 × 10 min in PBS, sections were treated for 30 min in PBS + 0.1% Triton X-100 (PBT) and incubated in 5% normal goat serum (NGS) for 30 min. Primary antibodies (rabbit anti-IR25a 1:500 (ref. 16), mouse anti-22C10 1:200, DSHB) were diluted in PBT with NGS and applied to slides placed horizontally in humidified chambers and left for 2 h at room temperature followed by incubation overnight at 4 °C. After washing for 3 × 10 min in PBT, slides were blocked in PBT with NGS for 30 min and incubated with secondary antibodies (rabbit AlexaFlour-594, 1:500, Mouse AlexaFluor-647 1:500, Invitrogen) diluted in PBT in the dark for 4 h at room temperature. Slides were washed 3 × 5 min in PBS and mounted in Vectashield before observation. Immunostaining of whole-mounted brains was performed as described in[36] with minor modifications. For LD experiments, flies were fixed on the fifth day of light entrainment. For temperature experiments, flies were first reared in LL and 25 °C for 3 days, and then transferred to a 25 °C:27 °C or 25 °C:29 °C temperature cycles for 7 days. Temperature cycles were rectangular and not ramped. Therefore, the conditions are consistent with those for behavioural analysis. For temperature entrainment in DD, flies were initially entrained to LD at 25 °C for 2 days followed by a 25 °C:27 °C temperature cycles in DD that was shifted 8 h in advance with respect to the previous LD cycle. Brains were dissected on day 6 of the temperature cycles at the indicated time points. Primary rat anti-TIM (1:1,000)[37], and secondary rat AlexaFluor-594 antibodies (Invitrogen, 1:500) were applied. Mounted brains were scanned using a Leica TCS SP5 confocal microscope. Quantification of TIM signals was performed as in[38] with minor modifications: Pixel intensity of stained neurons and background staining in each neuronal group was measured using ImageJ. Background signal was determined by taking the average signal of two surrounding fields of each neuronal group and was subtracted from the neuronal signal. For each group of clock neurons, at least 8 hemispheres from each genotype were checked and measured per time point. Data were normalized by setting the peak value to 1 and the ratio from each time point was then divided by the peak value.

Co-immunoprecipitation. Co-immunoprecipitation experiments were performed as described[14]. For each protein purification, 200–300 mg wet-weight of heads from *gmr-gal4* flies expressing *UAS-nocte-flag* (*FSNS* and *FSNH* transgenics were used in 2 independent experiments) or *gmr-gal4* alone (negative controls) were collected on dry ice and manually homogenized with a 2 ml Dounce homogenizer (Fisher) in 1 ml of extraction buffer (final protein concentration 5 mg ml^{-1} extraction buffer) containing 50 mM Tris, pH 7.5, 125 mM NaCl, 1.5 mM $MgCl_2$, 1 mM EDTA, 5% glycerol, 0.4% NP-40, and 0.1% Tween 20. To prevent degradation during the lengthy purification steps, 2× protease mini EDTA inhibitor mixture (Roche) was added at hourly intervals throughout the procedure. The homogenate was centrifuged at 10,000 r.p.m. for 15 min to isolate the soluble fraction used for pull-down. For the Flag pull-down procedure, EZview Red anti-Flag M2 affinity gel (Sigma) was used to bind the Flag-tagged bait and its bound partners. 50 μl prewashed 50% slurry was added to 1 ml soluble protein and incubated at 4 °C for 2 h on a rotary mixer. Non-binding material was removed by centrifugation (8,000*g* for 2 min) and the resin was washed three times in ice-cold extraction buffer. For checking the interaction with IR25a (Extended Data Fig. 1a), the washed resin was directly boiled with 5× SDS loading buffer followed by routine western blot. For elution the isolated protein complexes, Flag-tagged protein with any associating proteins, was incubated and eluted three times each with 50 μl (100 μg ml^{-1}) Flag peptide (Sigma) in extraction buffer for 30 min at 4 °C on a rotary mixer. The three eluates were combined and any residual resin was removed by centrifugation at 8,000*g* for 2 min. The following mass spectrometry peptide sequencing was performed by Cambridge Centre for Proteomics. Briefly, eluates from the tagged line and untagged control flies, were processed as described[14]. The only deviation from the method described was that peptides were applied to a 180 μm × 20 mm (5 μm particle size) C18 trap column (Waters UPLC Trap Symmetry) coupled to a nanoAcquity UPLC system (Waters) using 0.1% formic acid in water (buffer A) at a flow rate of 10 μl min^{-1}. Peptides were then separated on a 75 μm × 250 mm (1.7 μm particle size) reverse phase BEH C18 analytical nano-column (Waters) at a flow rate of 300 nl min^{-1} using a gradient of buffer A and buffer B (0.1% formic acid in acetonitrile). The HPLC system was directly coupled to a LTQ Orbitrap Velos (Thermo Scientific) with a New Objective nanospray ionisation source operated at a resolution of 60,000. Peptides were eluted with a linear gradient of 5–45% buffer B over 45 min or with a re-equilibration step, giving total running times of 60 min. The Orbitrap analyser survey scan was performed over a mass range of *m*/*z* 380–1,500 each of them triggering 10 MS2 LTQ acquisitions of the ten most intense ions exceeding 500 counts using a data dependent acquisition mode.

Western blot. For confirming the interaction between NOCTE and IR25a total head proteins were isolated from flies expressing IR25a, or IR25a and Flag-tagged NOCTE, under the control of *tim-gal4*. Boiled beads (after Co-IP) were loaded on SDS–PAGE gels, followed by standard western blot. Primary rabbit anti-IR25a 1:5,000 (ref. 16) and mouse anti-Flag M2 1:1,000 (Sigma), and secondary HRP-conjugated goat anti-rabbit IgG-HRP (1:10,000) and goat anti-mouse IgG-HRP (1:1,000) antibodies (Jackson) were used.

RNA isolation and RT–PCR. For RNA extractions, 30–50 flies were collected in 2 ml RNAlater (Ambion) and kept at 4 °C overnight, and 100 μl 0.1% PBST was added to help RNAlater penetration. Femurs from around 200 fly legs and 50 retinas were quickly dissected in cold RNAlater. Total RNA was extracted using an RNEasy kit (QIAGEN) according to the manufacturer's instructions. Total RNA was finally eluted in RNase-free water and stored at −80 °C. cDNA synthesis was performed with Reverse Transcription Reagents Kit (Applied Biosystems) in 10 μl reactions using 1 μg of total RNA according to the manufacturer's instructions. To verify mRNA expression level of *IR25a* and *nocte* in fly femur ChO, dilutions of cDNA were used for PCR with the following primers: *rp49* and *nocte* (ref. 9), *IR25a* (ref. 39), followed by DNA electrophoresis on 2% agarose gels to visualize the PCR products. To test the efficiency of *IR25a* RNAi, 20 heads of 5–10-day-old flies were dissected and RNA was extracted and reverse transcribed as described above. Taqman probes for *IR25a* (catalogue number 4351372, ThermoFisher) and *RPL32* (catalogue number 4448489, ThermoFisher) were applied to determine the amount of mRNA. For determining *IR25a* mRNA levels in different tissues, body parts were dissected in RNA later (Ambion). 1 μg RNA was used for cDNA synthesis. Real-time assays were performed using an ABI GeneAMP PCR system 9700 using the standard program, and C_t (threshold cycle) values were applied to

determine the amount of RNA in each genotype. The relative concentrations were calculated using the $2^{-\Delta\Delta C_t}$ method, and *RPL32* was used as control.

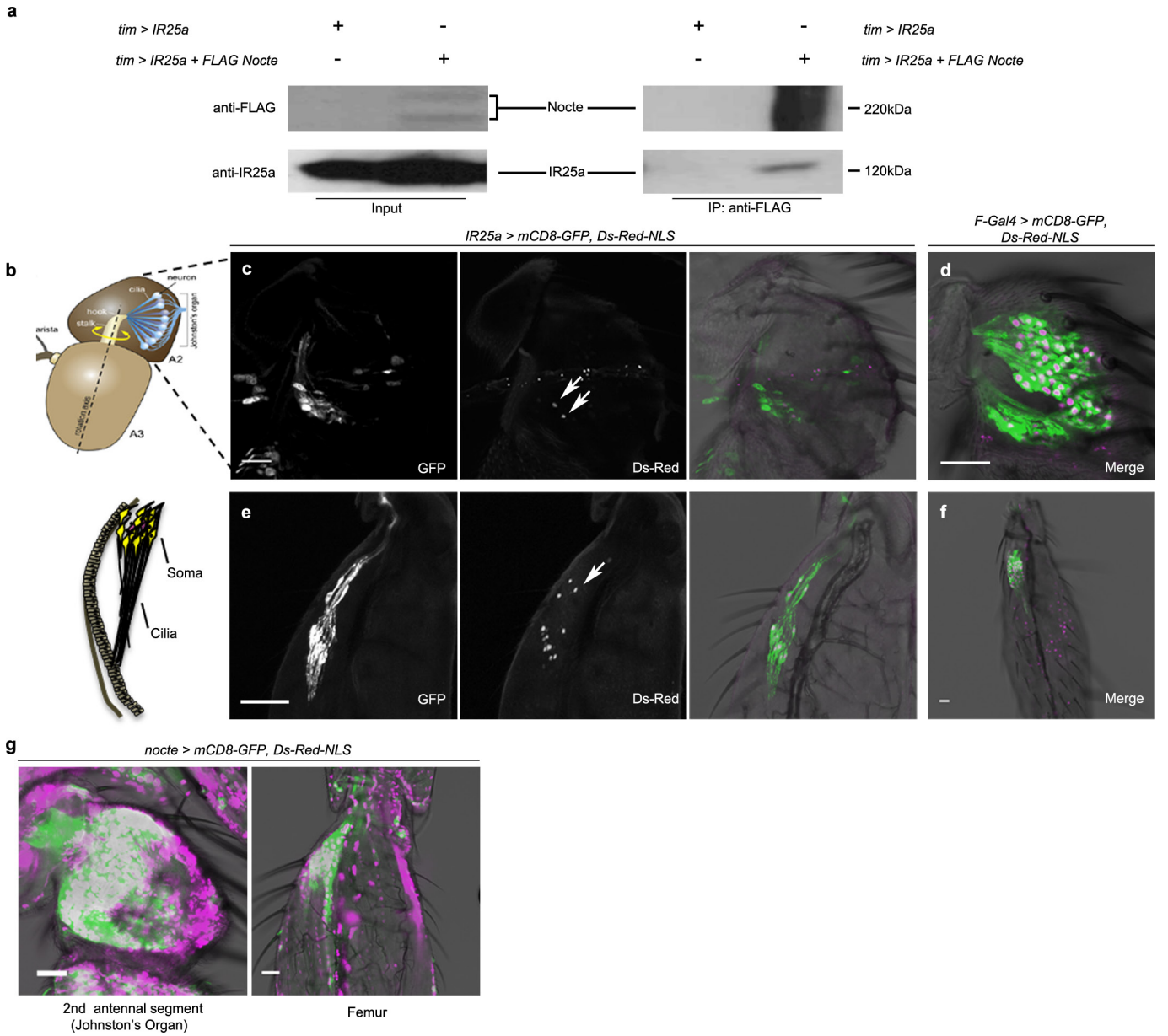
Behavioural analysis. Analysis of locomotor activity of 4–5-day-old male flies was performed using the *Drosophila* Activity Monitor System (DAM, Trikinetics). The DAM monitors, as well as an environmental monitor (Trikinetics), were located inside a light- and temperature-controlled incubator where the fly's activity was monitored for a few weeks depending on different experimental conditions. Plotting of behavioural activity and period calculations were performed using a signal-processing tool-box[40] implemented in Matlab (MathWorks). In order to quantify behaviour during temperature cycles, an updated Histogram version based on Excel (Office, Microsoft) was applied[38]. Briefly, the activity from the last two days of each temperature cycle was plotted in Excel in 30-min bins and an 'entrainment index' (EI = ratio of activity occurring during the 6 h window covering the main activity peak of the positive controls over the activity during the entire warm phase) was calculated. To distinguish the clock-controlled behavioural peaks from temperature response peaks, a simple smoothing filter was applied for the four activity bins during the 2 h following each temperature transition[38]. The filtered data was used for calculation, whereas the raw activity data are plotted in the histograms. The entrainment index values plotted in all histograms represent the average entrainment index from the temperature cycles before and after the shift except for 18 °C:20 °C and 21 °C:23 °C, where the entrainment index was generated from the temperature cycles after the shift. To calculate the phase of the main activity peaks after DD and temperature cycles involving genotypes that did not show clear activity peaks during entrainment, we employed circular phase plot analysis as previously described[40,41]. In brief, the mean activity phase of the three consecutive days after release into constant conditions was determined for each fly of the two genotypes to be compared. An average 'vector' indicating phase coherence (length) and mean peak phase (direction) is calculated for each genotype and the two vectors are compared by an *F*-statistic (Watson–Williams–Stevens test). The difference in direction is plotted in hours (h) and the mean peak phase of the controls (negative or positive controls, depending on the experiment) was set to zero.

Electrophysiology. *Extracellular leg nerve recording.* The question of whether ChO in the legs would respond to temperature changes was examined using extracellular recordings in restrained intact leg preparations. Canton S w^+ flies were used as a control and $IR25a^{-/-}$ mutants were used to test if IR25a is involved in this process. In order to minimize locomotory artefacts, flies were decapitated and all legs but the left hind leg amputated. Flies were mounted ventral side up and pinned down in a Sylgard (Dow Corning, USA) coated recording chamber so that the left hind leg was orientated perpendicular to the body but not immobilized (Fig. 4a). Therefore the ChO could be stimulated *in vivo* by moving the tibia with a fine needle. A tungsten wire electrode, sharpened to a fine point, was inserted through the cuticle in the thorax served as a reference electrode and a similar recording electrode was placed in the coxa of the remaining leg. The final position of the recording electrode was determined by monitoring the signal it was recording and then manually extending the tibia until a response in sensory units was seen. The signal was amplified using a BioAmp extracellular amplifier, filtered (low 5 kHz, high 10 Hz), digitized (sampling frequency 10 kHz) with a PowerLab 2/20 and recorded using LabChart 7 (ADInstruments, Bella Vista, Australia).

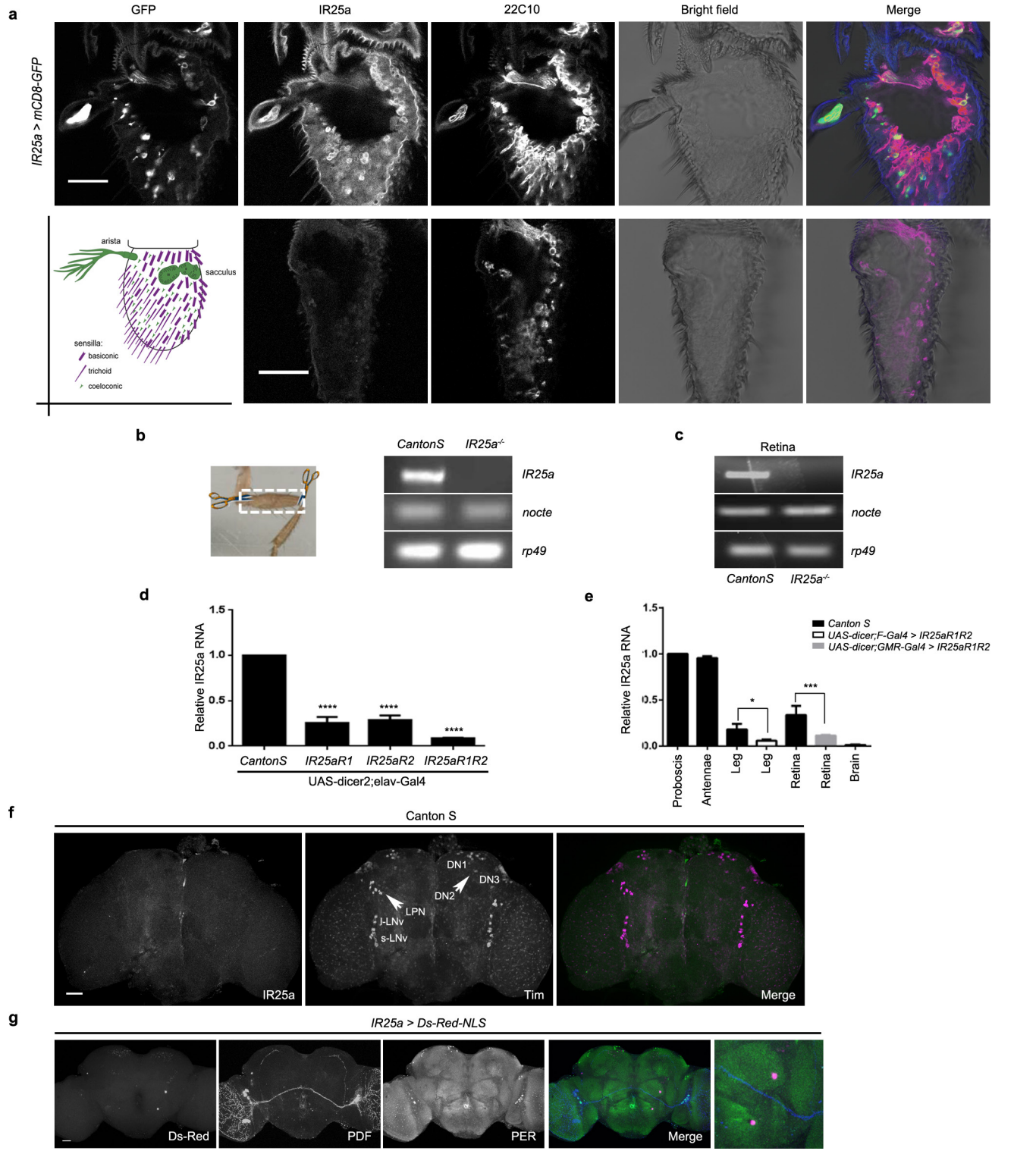
Whole-cell recordings. Different genotypes were used for each group and the data pooled as there were no differences between them: control, *Pdf-RFP* and *Pdf-gal4; UAS-mCherry*; IR25a, *Pdf-gal4/UAS-IR25a; Pdf-RFP* and *Pdf-gal4/UAS-mCherry-IR25a*; TrpA1, *Pdf-gal4/UAS-TrpA1; Pdf-RFP* and *Gal1118-gal4/UAS-TrpA1/UAS-mCherry-IR8a*; *Pdf-gal4/UAS-IR8a; Pdf-RFP* and *Pdf-gal4/UAS-mCherry-IR8a*. Experiments were performed under red light illumination and light exposure during dissection was kept to a minimum. For visualization of the l-LNv we used RFP-tagged constructs and a 555 nm LED light source in order to not activate cryptochrome. Adult flies raised in 12 h:12 h LD at 25 °C, were collected ~3–5 days post eclosion between ZT13 and ZT16, decapitated and brains dissected in extracellular saline solution containing (in mM): 101 NaCl, 1 $CaCl_2$, 4 $MgCl_2$, 3 KCl, 5 glucose, 1.25 NaH_2PO_4, 20.7 $NaHCO_3$, pH adjusted to 7.2. The brains were transferred for 5–10 min to saline containing 20 U per ml papain with 1 mM L-cysteine to digest the ganglion sheath. After removal of the photoreceptors, air sacks and trachea, a small incision was made over the position of the l-LNv neurons in order to give easier access for the recording electrodes. Brains were placed ventral side up in the recording chamber, secured using a custom-made anchor and during recordings continuously perfused with aerated (95% O_2, 5% CO_2) saline solution. l-LNv neurons were identified on the basis of their fluorescence, size and position. A single recording was performed from one l-LNv per brain. Whole-cell current clamp recordings were performed using glass electrodes with 10–20 MΩ resistance filled with intracellular solution (in mM: 102 K-gluconate, 17 NaCl, 0.94 EGTA, 8.5 HEPES, 0.085 $CaCl_2$, 1.7 $MgCl_2$, pH 7.2) and an Axon MultiClamp 700B amplifier, digitized with an Axon DigiData 1440A (sampling rate: 20 kHz; filter: Bessel 10 kHz) and recorded using pClamp 10 (Molecular Devices, USA). A cell was included in the analysis if the access resistance was less than 70 MΩ and the leak current in response to a −40 mV pulse less than −100 pA. All chemicals were purchased from Sigma (Poole, UK). The liquid junction potential was calculated as 13 mV and was subtracted from all the membrane voltages. Resting membrane potential (*V*m) was measured after stabilizing for 2–3 min. Membrane input resistance (*R*in) was calculated by injecting hyperpolarizing current steps and measuring the resulting changes in voltage. Spike frequency was manually measured using 10 s bins for each degree of temperature. To test the effect of elevated temperature, the recording chamber and the perfusion influx were gradually heated from 18 °C to 30 °C within 5–10 min and cooled back to 18 °C within 10–15 min using a Peltier heating system (ALA Scientific Instruments, USA) and TC-10 controller (npi, Tamm, Germany). The temperature coefficient Q10 was calculated by dividing the firing rate at 30 °C by the rate at 20 °C. To check whether l-LNvs can also sense small temperature changes of 2–3 °C, neurons were recorded as before, the temperature increased to around 24.5 °C held for 3 min, then increased to 27.5 °C, held again for 3 min, cooled back down to 24.5 °C and recorded for a further 3 min. During the whole period the instantaneous spiking frequency was monitored. All values are given as mean and s.e.m. and a *t*-test and ANOVA (followed by Tukey test) were used to calculate significant differences.

Data reporting. No statistical methods were used to predetermine sample size. The investigators were not blinded to allocation during experiments and outcome assessment.

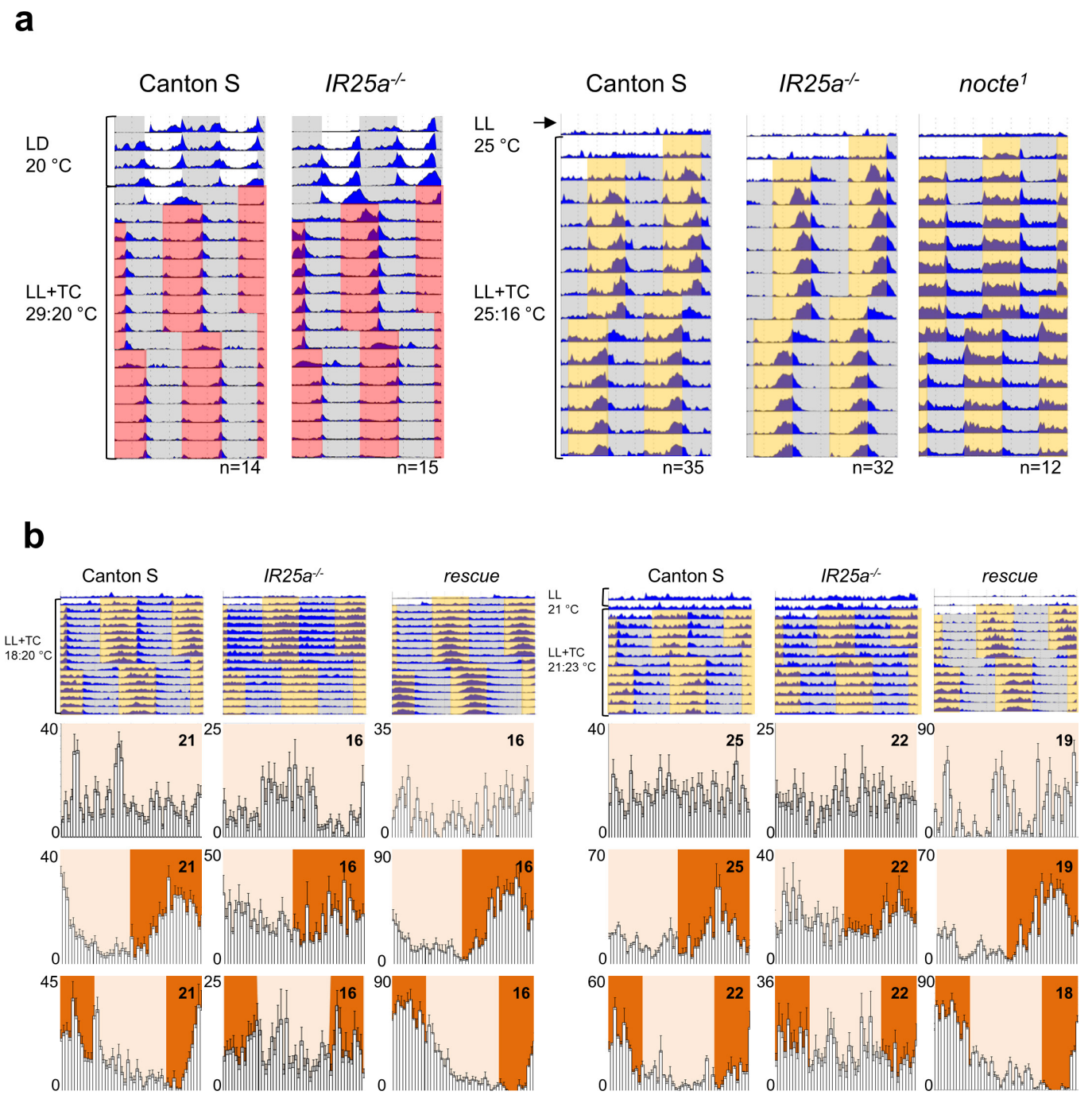
24. Venken, K. J. *et al.* Versatile P[acman] BAC libraries for transgenesis studies in *Drosophila melanogaster*. *Nature Methods* **6,** 431–434 (2009).
25. Kaneko, M. & Hall, J. C. Neuroanatomy of cells expressing clock genes in *Drosophila*: transgenic manipulation of the *period* and *timeless* genes to mark the perikarya of circadian pacemaker neurons and their projections. *J. Comp. Neurol.* **422,** 66–94 (2000).
26. Gummadova, J. O., Coutts, G. A. & Glossop, N. R. Analysis of the *Drosophila Clock* promoter reveals heterogeneity in expression between subgroups of central oscillator cells and identifies a novel enhancer region. *J. Biol. Rhythms* **24,** 353–367 (2009).
27. Kim, J. *et al.* A TRPV family ion channel required for hearing in *Drosophila*. *Nature* **424,** 81–84 (2003).
28. Park, J. H. & Hall, J. C. Isolation and chronobiological analysis of a neuropeptide pigment-dispersing factor gene in *Drosophila melanogaster*. *J. Biol. Rhythms* **13,** 219–228 (1998).
29. Liu, L. *et al.* *Drosophila* hygrosensation requires the TRP channels water witch and nanchung. *Nature* **450,** 294–298 (2007).
30. Chung, Y. D., Zhu, J., Han, Y. & Kernan, M. J. *nompA* encodes a PNS-specific, ZP domain protein required to connect mechanosensory dendrites to sensory structures. *Neuron* **29,** 415–428 (2001).
31. Ruben, M., Drapeau, M. D., Mizrak, D. & Blau, J. A mechanism for circadian control of pacemaker neuron excitability. *J. Biol. Rhythms* **27,** 353–364 (2012).
32. Konopka, R. J. & Benzer, S. Clock mutants of *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA* **68,** 2112–2116 (1971).
33. Sweeney, S. T., Broadie, K., Keane, J., Niemann, H. & O'Kane, C. J. Targeted expression of tetanus toxin light chain in *Drosophila* specifically eliminates synaptic transmission and causes behavioral defects. *Neuron* **14,** 341–351 (1995).
34. Zhang, Y., Liu, Y., Bilodeau-Wentworth, D., Hardin, P. E. & Emery, P. Light and temperature control the contribution of specific DN1 neurons to *Drosophila* circadian behavior. *Curr. Biol.* **20,** 600–605 (2010).
35. Saina, M. & Benton, R. Visualizing olfactory receptor expression and localization in *Drosophila*. *Methods Mol. Biol.* **1003,** 211–228 (2013).
36. Yoshii, T., Todo, T., Wülbeck, C., Stanewsky, R. & Helfrich-Förster, C. Cryptochrome is present in the compound eyes and a subset of *Drosophila*'s clock neurons. *J. Comp. Neurol.* **508,** 952–966 (2008).
37. Rush, B. L., Murad, A., Emery, P. & Giebultowicz, J. M. Ectopic CRYPTOCHROME renders TIM light sensitive in the *Drosophila* ovary. *J. Biol. Rhythms* **21,** 272–278 (2006).
38. Gentile, C., Sehadova, H., Simoni, A., Chen, C & Stanewsky, R. Cryptochrome antagonizes synchronization of *Drosophila*'s circadian clock to temperature cycles. *Curr. Biol.* **23,** 185–195 (2013).
39. Croset, V. *et al.* Ancient protostome origin of chemosensory ionotropic glutamate receptors and the evolution of insect taste and olfaction. *PLoS Genet.* **6,** e1001064 (2010).
40. Levine, J. D., Funes, P., Dowse, H. B. & Hall, J. C. Signal analysis of behavioral and molecular cycles. *BMC Neurosci.* **3,** 1 (2002).
41. Simoni, A. *et al.* A mechanosensory pathway to the *Drosophila* circadian clock. *Science* **343,** 525–528 (2014).
42. Wilson, R. I. & Corey, D. P. The force be with you: a mechanoreceptor channel in proprioception and touch. *Neuron* **67,** 349–351 (2010).
43. Stanewsky, R. *et al.* Temporal and spatial expression patterns of transgenes containing increasing amounts of the *Drosophila* clock gene *period* and a *lacZ* reporter: mapping elements of the PER protein involved in circadian cycling. *J. Neurosci.* **17,** 676–696 (1997).



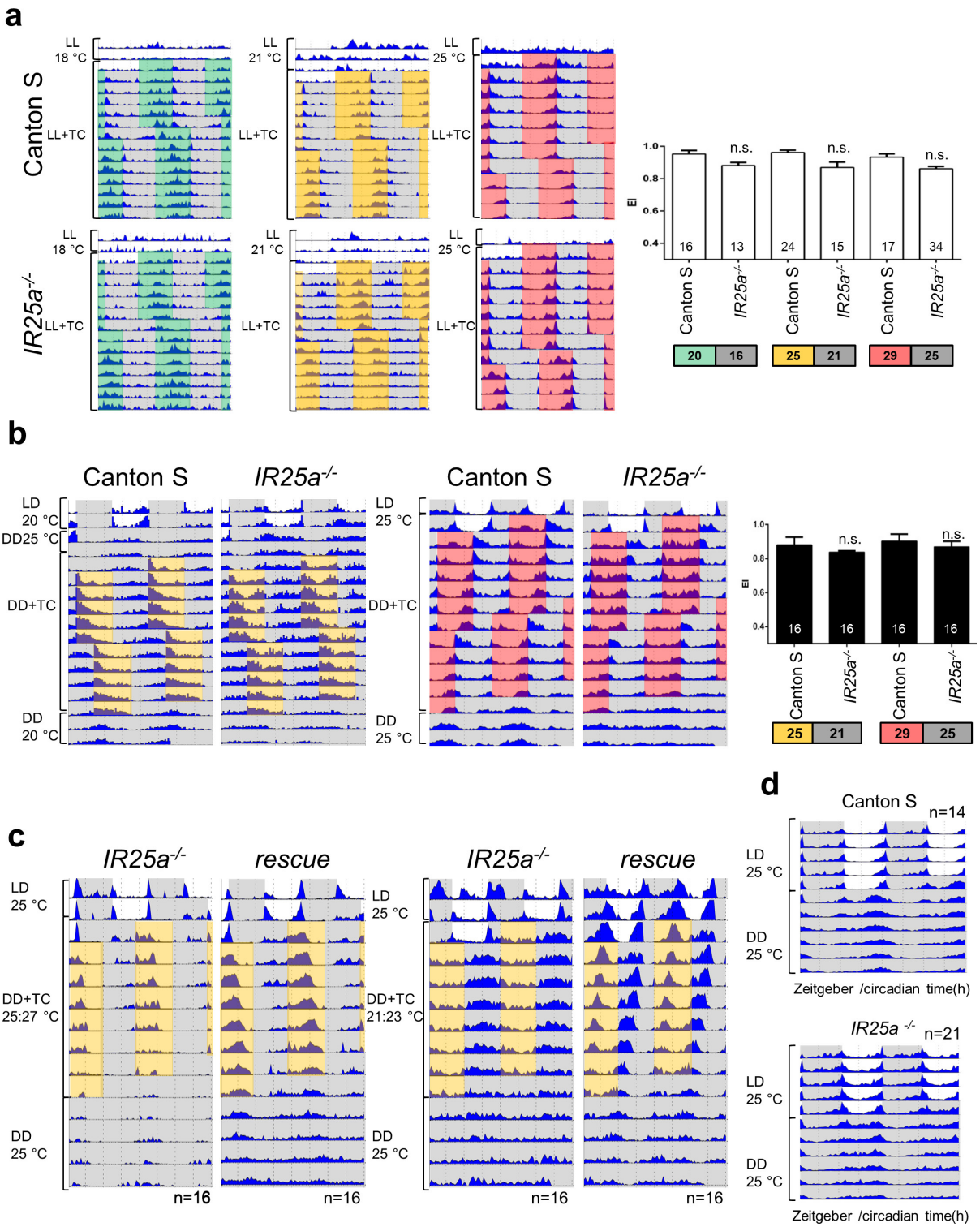
Extended Data Figure 1 | IR25a and Nocte physically interact *in vivo* and are expressed in femur and antennal ChO neurons. a, *In vivo* co-immunoprecipitation experiments using protein extracts from fly heads. Head lysates were immunoprecipitated using anti-Flag antibody. The immunoprecipitates were examined by western blotting using anti-Flag and anti-IR25a antibody. Input represents 30% of cell lysates used in the pull-down experiment. The genotypes of the flies used were: *tim > IR25a*: *UAS-GFP/UAS-IR25a; tim-gal4:67/+*. *tim > IR25a+FLAG-NOCTE*: *UAS-FSNH/UAS-IR25a; tim-gal4:67/+*. The bracket indicates that NOCTE–Flag runs as a double band on western blots. For uncropped gel images, see Supplementary Fig. 1. **b**, Overview of the antennal and femur ChO adapted from refs 13, 42. **c**, **d**, Labelling of the JO neurons by *IR25a-gal4* (**c**) and *F-gal4* (**d**) driven membrane bound mCD8–GFP and nuclear-localized DsRed expression. Note that *IR25a* is expressed in only a subset of JO neurons. **e**, **f**, Same flies as in **c**, **d** analysed for *IR25a* expression in the femoral ChO. Again, only subsets of the ChO neurons express *IR25a*. Arrows in **c**, **e** point to ChO neuron nuclei. **g**, Labelling of the ChO neurons in JO and femur by *nocte-gal4*-driven membrane-bound GFP and nuclear DsRed expression[9]. Scale bar, 20 μm.



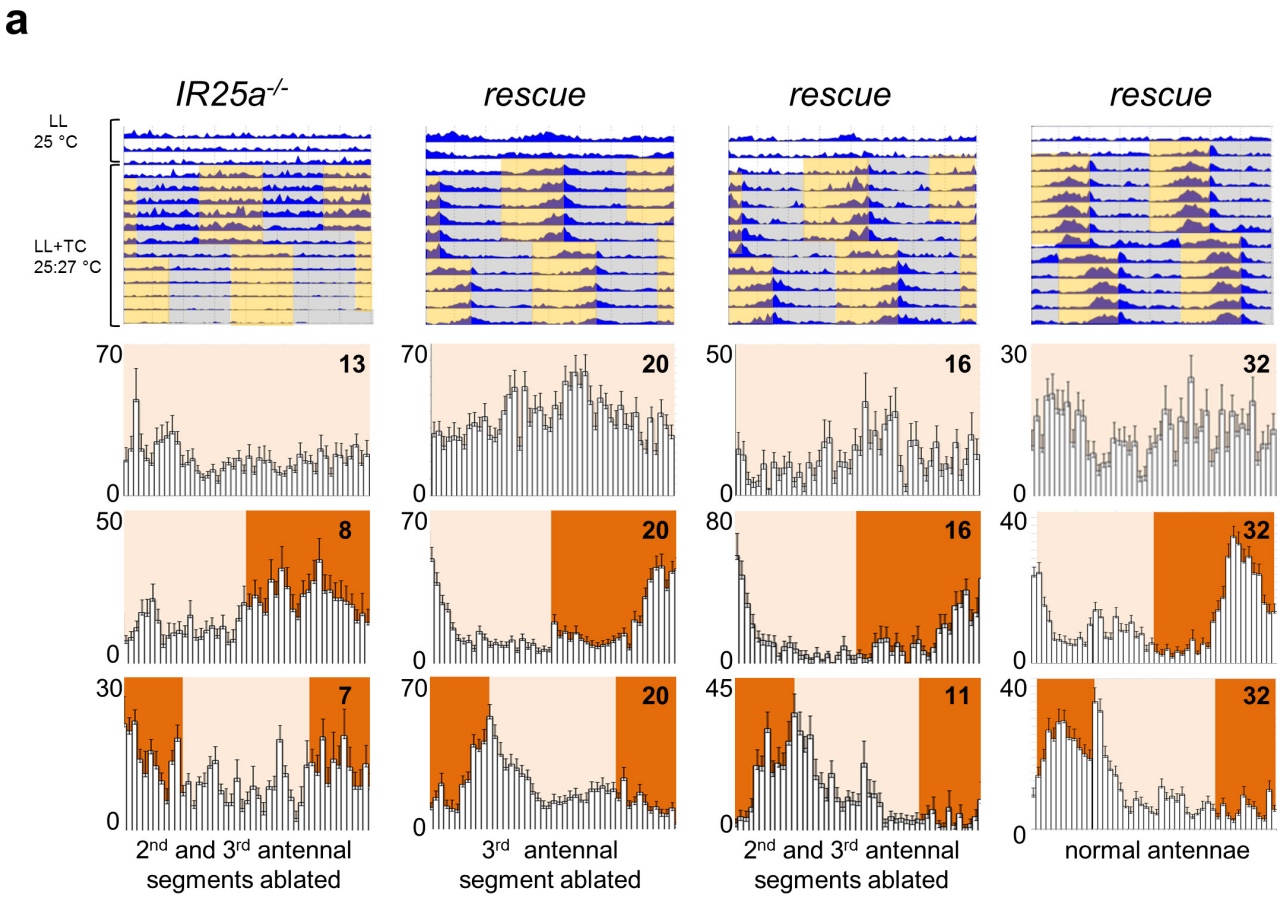
Extended Data Figure 2 | Spatial and quantitative *IR25a* mRNA and protein expression in CNS and PNS tissues and efficiency of RNAi-mediated knockdown. a, Analysis of *IR25a-gal4* and IR25a in the third segment of the antenna reveals expression in coeloconic sensilla[16]. Schematic adapted from[15]. **b**, **c**, Determination of *IR25a* and *nocte* mRNA levels in femur and retinal tissues by semiquantitative RT–PCR; *rp49* was used as control. For uncropped gel data, see Supplementary Fig. 1. **d**, **e**, qPCR analysis of *IR25a* mRNA levels in whole heads (**d**), or dissected body parts (as indicated) (**e**) from flies of the genotypes indicated. Pan-neuronal *elav-gal4* knockdown (**d**) decreased *IR25a* mRNA >75% or >90%, using one or two different RNAi lines combined, respectively. $****P < 0.0001$, $***P < 0.001$, $*P < 0.05$, one-way ANOVA followed by Bonferroni correction. **f**, IR25a is not expressed in the central brain and clock neurons. Left, IR25a immunolabelling of a Canton S brain reveals no signals. Middle, same brain labelled with anti-TIM reveals expression in clock neurons. Right, merge. Brains were dissected in LD at ZT20. Scale bar, 10 μm. **g**, *IR25a-gal4* is not expressed in clock neurons and largely absent from the brain. Left, nuclear DsRed driven by *IR25a-gal4*. Second from left, anti-PDF staining showing LNv and their projections. Middle, anti-PER (diluted 1:5,000)[43] showing all clock neurons. Second from right, merge, showing two *IR25a-gal4* positive cell in the antennal lobe, not co-localized with any of the clock neurons. These cells were observed in 4/8 hemispheres and always on the same side of the brain. Right, magnified view of circled area in the merged image. Scale bar, 30 μm.

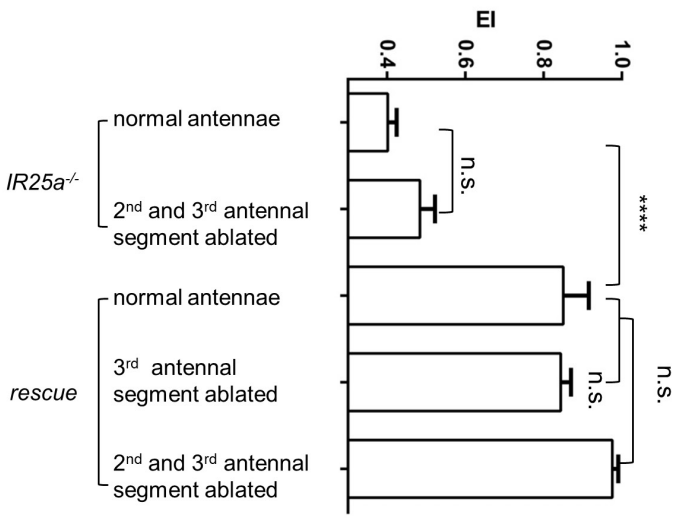


Extended Data Figure 3 | *IR25a* is required for temperature synchronization to low-amplitude temperature cycles but not for high-amplitude temperature cycles. a, Canton S, $IR25a^{-/-}$, and *nocte*[1] flies were exposed to LD at 20 °C for 5 days (left) or LL at 25 °C for 2 days (right), followed by exposure to a 12 h:12 h 20 °C:29 °C (left) or 16 °C:25 °C (right) temperature cycles in LL, which after 6–7 days was delayed or advanced by 6 h, respectively. Warmer temperature indicated by red and orange shading, respectively. **b**, Actograms and daily averages of Canton S and $IR25a^{-/-}$, and $IR25a^{-/-}$ flies containing a genomic *IR25a* rescue construct (rescue) exposed to 18 °C:20 °C temperature cycles in LL (left) and 21 °C:23 °C temperature cycles in LL (right). Warm phase in actograms indicated by orange shading. Histogram colour coding as in Fig. 2. For quantification see Fig. 2e.

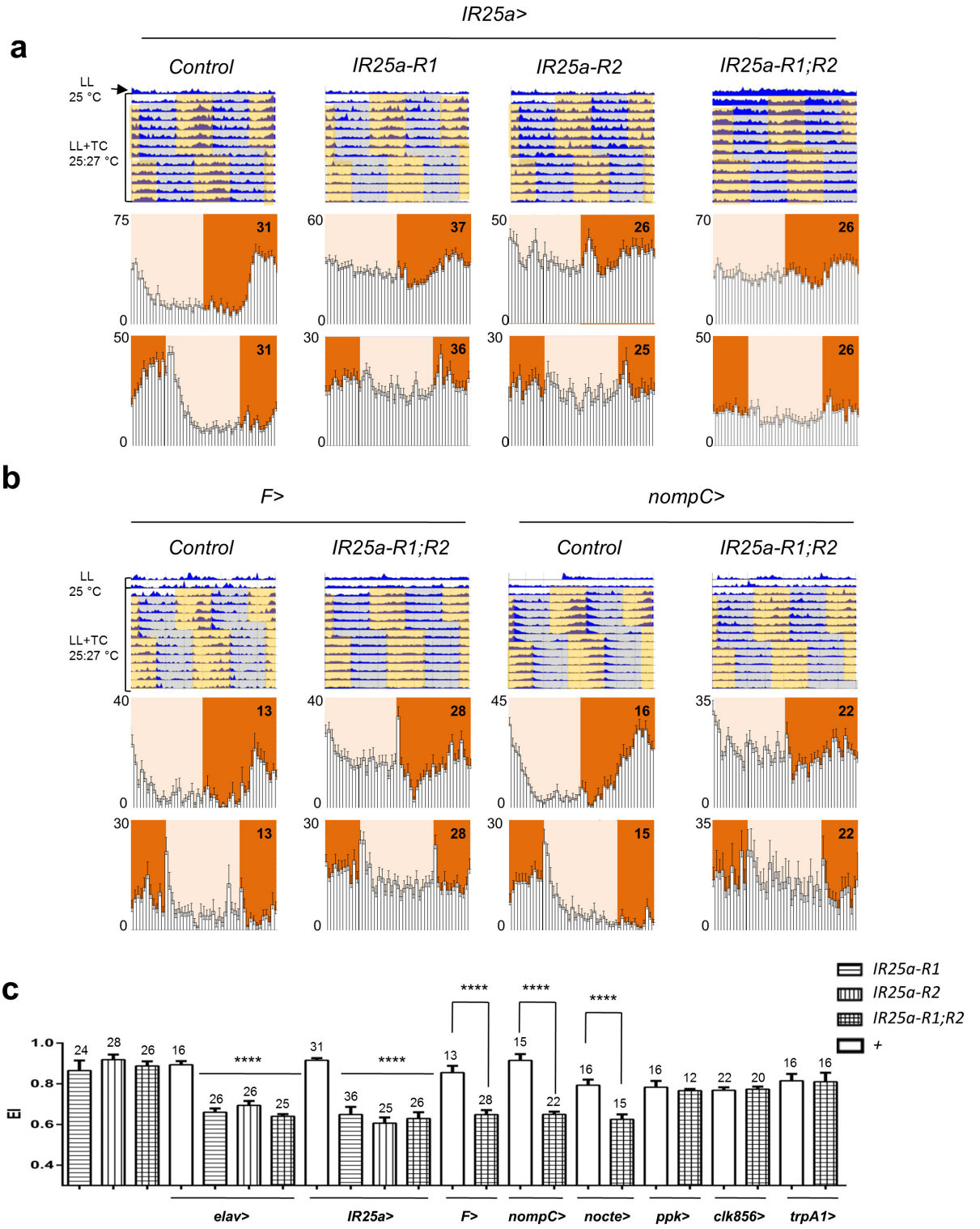


Extended Data Figure 4 | *IR25a* is not required for temperature synchronization to high- but to low-amplitude temperature cycles and *IR25a*$^{-/-}$ flies show normal LD and DD behaviour. a, Canton S and *IR25a*$^{-/-}$ flies were exposed to LL at 25 °C for 2–3 days, followed by exposure to high-amplitude 12 h:12 h temperature cycles in LL, which after 5–6 days were delayed by 6 h. Double plotted average actograms depicting the daily activity levels and environmental conditions during the entire experiment are shown. Actual temperatures are colour coded and indicated below the entrainment index calculations. Numbers (*n*) indicated in the bars. **b**, As in **a** but flies were initially kept in LD and DD for 2 days each (left) or LD (right), before being exposed to two phase delayed (left) or advanced (right) temperature cycles in DD at the temperatures indicated. n.s., not significant. **c**, Behaviour of *IR25a*$^{-/-}$ and rescue flies during DD and 25 °C:27 °C temperature cycles with 8 h delay and during DD and 21 °C:23 °C temperature cycles with a 8 h advance compared to the previous LD cycle (at 25 °C). Warm phase is indicated by orange shading. **d**, Canton S and *IR25a*$^{-/-}$ flies during LD and DD conditions at 25 °C (see Extended Data Table 2 for period calculations).

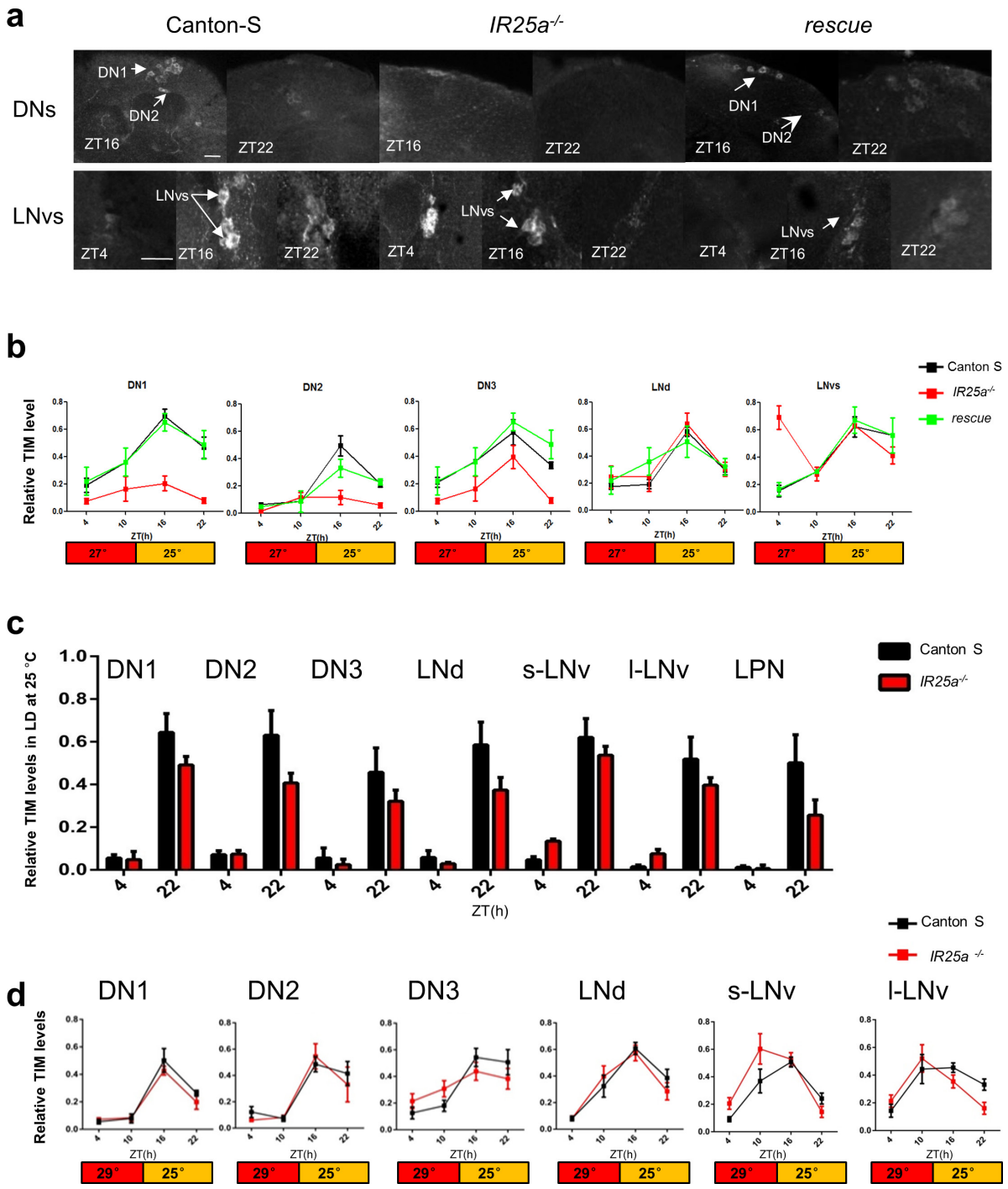




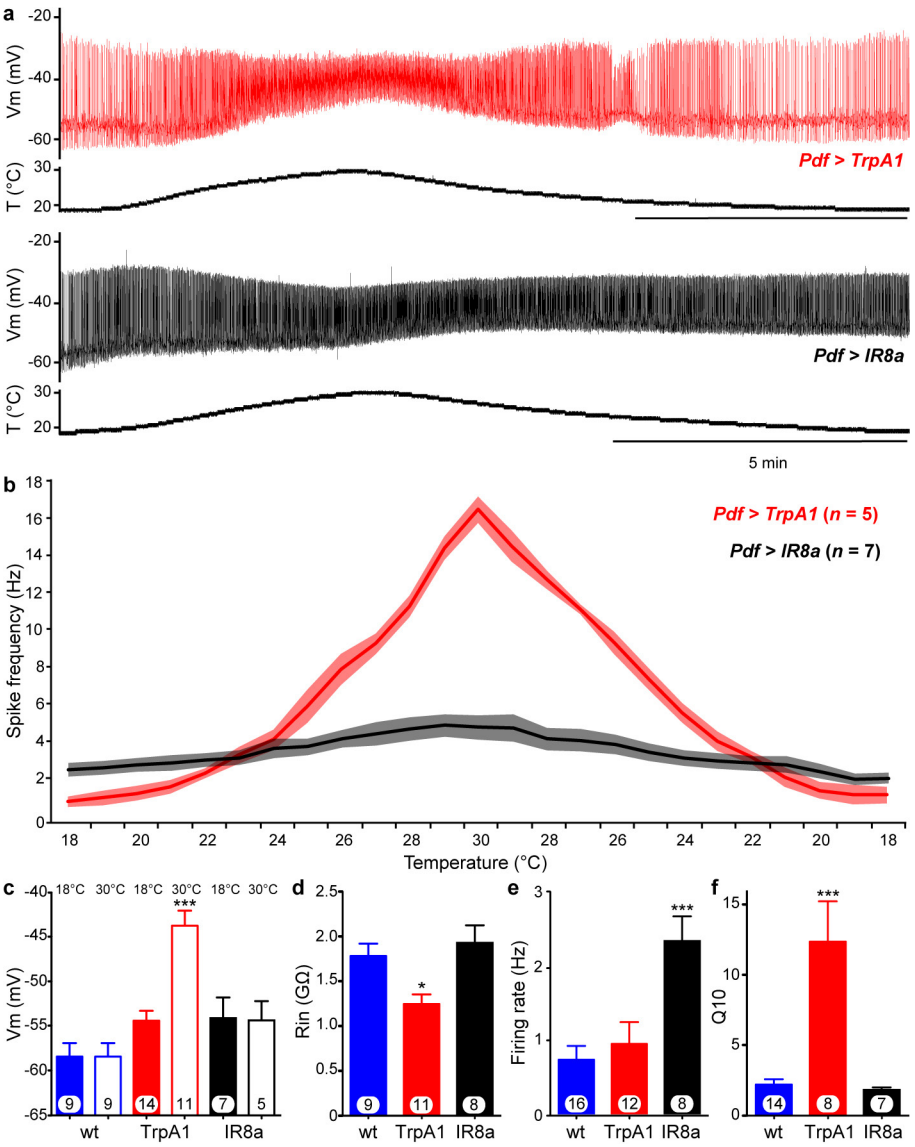
Extended Data Figure 5 | Antennal *IR25a* expression is not necessary for synchronization of locomotor activity rhythms to temperature cycles. Ablation of antennae as indicated. **a**, $IR25a^{-/-}$ and rescue flies were exposed to the same condition used in Fig. 2. Actograms and daily averages as described before. **b**, Quantification of behaviour as described in Fig. 2. The data of $IR25a^{-/-}$ with normal antennae was taken from Fig. 2a. n.s., not significant.



Extended Data Figure 6 | Knocking down *IR25a* expression via RNA interference disrupts synchronization of locomotor activity rhythms to temperature cycles (25 °C:27 °C in LL). **a**, **b**, Behaviour of flies with spatially restricted *IR25a* knockdown mediated by *IR25a-gal4* (**a**), *ChO* specific *F-gal4,* and *nompC-gal4* (**b**) driven *IR25-RNAi* expression, respectively. Control flies are *UAS-dicer2/Y; IR25-gal4/+* (**a**), and *UAS-dicer2/Y;+/+; F-gal4/+* or *UAS-dicer2/Y;+/+; nompC-gal4/+* (**b**). Test flies carry the same transgenes, but in addition one or two copies of the *IR25a-RNAi* line indicated. Actograms and daily averages as described for Fig. 2a. **c**, Progeny of the respective *UAS-IR25a-RNAi* lines crossed to *y w* (left three columns) and flies from (**a**, **b**) and the other *gal4* drivers indicated, were exposed to the same LL and temperature cycle conditions used in Fig. 2a. As controls, *UAS-dicer, gal4* driver lines were crossed to *y w* and F1 males containing *UAS-dicer/Y* and the respective *gal4/+* were tested. Numbers of analysed individuals (*n*) are indicated above each column. Entrainment was quantified as in Fig. 2c. ****$P < 0.0001$, one-way ANOVA followed by Bonferroni correction.



Extended Data Figure 7 | Rescue of TIM oscillations in clock neurons during low-amplitude temperature cycles and normal TIM oscillations during LD and high-amplitude temperature cycles. a, **b**, TIM levels in clock neurons during LL 25 °C:27 °C temperature cycles at the indicated time points (ZT) in the genotypes indicated. At least 8 brain hemispheres per time point were analysed for each genotype. Scale bars, 10 μm. Data in **b** are mean ± s.e.m. **c**, Quantification of TIM levels in clock neurons during LD (25 °C) in Canton S and *IR25a*$^{-/-}$ mutant brains. **d**, TIM oscillations in different clock-neuronal groups in *IR25a*$^{-/-}$ are restored in 25 °C:29 °C temperature cycles in LL. At least 8 brain hemispheres per time point were analysed for each genotype and condition. Error bars indicate s.e.m.



Extended Data Figure 8 | Ectopic expression and heat responses of TRPA1 and IR8a in l-LNv clock neurons. **a**, Whole-cell current clamp recordings of *Pdf-gal4/UAS-TrpA1; Pdf-RFP* (top trace, red) and *Pdf-gal4/UAS-IR8a-RFP* (bottom trace, black) brains exposed to a temperature ramp from 18 °C to 30 °C and back to 18 °C. Note the additional depolarization of the *TrpA1* neuron at higher temperatures. **b**, Compared to control (Fig. 4), recordings from *TrpA1*-expressing neurons show a large increase in firing rate with temperature which the IR8a expressing neurons do not. **c**, **d**, In comparison to control neurons (data taken from Fig. 4) the membrane potential of *TrpA1* expressing neurons is more positive at 30 °C (open bars) and the input resistance is also significantly reduced in *TrpA1* at 18 °C. **e**, **f**, The firing rate at 18 °C is higher for IR8a neurons but only the Q10 of *TrpA1* is different to control. Bars are means and whiskers s.e.m., *n* indicated in bars, $*P < 0.05$, $***P < 0.001$, ANOVA followed by Tukey test.

Extended Data Table 1 | Mass spectrometry data from fly heads from three different genotypes

		UNIQUE PEPTIDES			% SEQ. COVERAGE			MASCOT SCORE		EMPAI SCORE		PEPTIDE SEQUENCES	
FLYBASE ID GENE	DESCRIPTION	FSNS	FSNH	NEG	FSNS	FSNH	NEG	FSNS	FSNH	FSNS	FSNH	FSNS	FSNH
FBpp0298823 nocte	**no circadian temperature entrainment, isoform C**	**22**	**17**	**4**	**10.87**	**8.79**	**1.99**	**1399.5**	**713**	**0.30**	**0.20**	**LSASTTSWQR,LGYEEYGK,GASVGGSSGYGR,SGYGR,SISGGYVQR,QPVGTGSAGGSGSGGSGR,QDDIDFTK,QAQALPR,GYAGSSGGSSVGSGSSGGSSVSGSSYR,GGVSGGGGGVSGGAQANGSGGSGR,NASDWGSSR,TLTTDPMPTQILR,TSESETDLDKTK,QQQQQQQQLPR,SPLSADMSLGLAK,KLQELEMK,SASASSAFDSNSR,EQAAAAAVAAQR,LGFSFGDDPTTPLK,FTALDINR,KIESCAVVGGEK,SASPAVVGSGSFR**	**LGYEEYGK,GASVGGSSGYGR,QPVGTGSAGGSGSGGSGR,QAQALPR,GYAGSSGGSSVSGSSYR,RQPVGTGSAGGSGSGGSGR,GGVSGGGGGVSGGAQANAGQGR,TLTTDPMPTQILR,TSESETDLDKTK,QQQQQQQQLPR,SPLSADMSLGLAK,KLQELEMK,SASASSAFDSNSR,EQAAAAAVAAQR,LGFSFGDDPTTPLK,FTALDINR,SASPAVVGSGSFR**
FBpp0079064 ninaC	neither inactivation nor afterpotential C, isoform B	20	8	2	15.66	5.8	1.27	447	102.5	0.25	0.09	AILMLVNAGTPVNNDSTR,MYPEDLAALENPVDENIIESLR,AMFQIIR,LCDFGLSR,TLYKEPELFVDR,SSLDESIMLMFTNQLTK,DAVASTLYSR,YQFLAFDFDEPVEMTK,LPFDEFLR,TALDNLLTKPDGLFYIIDDASR,YAEVENTDIVSR,YYNDEFLAR,MGESDNIYNQGYFR,SDIAEMLELSR,AFTDINR,QYTTEEAR,ADLEYKPR,SCQDQDLIMDR,LVDFIINR,AAIELNR	LPFDEFLR,TALDNLLTKPDGLFYIIDDASR,TLYKEPELFVDR,SDIAEMLELSR,QYTTEEAR,SCQDQDLIMDR,LVDFIINR,AAIELNR
FBpp0072672 alpha-Spec	alpha spectrin, isoform A	46	7	2	21.61	3.73	0.95	1317.5	81.5	0.36	0.04	VSTLGAEAQR,LLDSYDLQR,QNQINSQYDNLLALAR,DLIGVQNLIK,FATDDSYLDPTNLNGK,QEAFLANEDLGDSLDSVEALIK,LMDVSNLGVPEIEQR,VTEVNQLADK,DLTGVQNLK,CNSIEEIR,DQPFASDDIR,DLASVQALQR,QQETPVVDITGK,LLAMQEQFR,LNEACQQQQFNR,FIESGHFDADNIR,DADETVAWIAEK,QGFVPAAYIK,MQEIVVLWETLVQASDK,IQSVLAMGGNLIDK,RAALQEK,LLNVISSGENMLK,FDDFNDDLK,QAEIANYWQSLTTK,YAALAAPMGER,DVVLSSDDYGR,DVAGAEALLER,MQEIVVLWETLVQASDKK,LLVGSDDYGR,LGDEQTLQQFSR,ILYEQCMDLQLFYR,LQAASEESYRDPTNLQAK,DLEDEAAWIR,QLLEDSNR,EKEPIAASTNR,QLDETANR,ALDIFATK,ETENVQSYEEIENAFR,AIISADELAK,ALAALDQK,ILETVEDIQER,ALNQAWAELK,NKEGNLSAR,IQTQMQDLNEK,LIDGQHYAADDVAQR,DADEIENWIAEK	QEAFLANEDLGDSLDSVEALIK,LLNVISSGENMLK,ILETVEDIQER,ALNQAWAELK,YAALAAPMGER,IQTQMQDLNEK,LNEACQQQQFNR
FBpp0074228 beta-Spec	beta spectrin, isoform A	21	1	0	10.82	0.52	0	472.5	33	0.14	0.01	LMTLSNSR,QVLLSQSLDQQLFNR,LVSQDNFGFDLAAVEAAAK,YATSQDESYR,VVGIAMENDK,LTTFELK,LLEVLSGER,QFSGDDISSGDAVAVK,LTSEGSTSDDPQYMFLR,IVSSEDVGRDEANVQSLLK,GAMLFDANR,MLDTMTPGKDIEDVEIMK,LNEALQLFMLSR,LEAIDNR,AQSDSTAVAASR,VAVVNQLAR,RQELAEEAER,LKQETVQGK,ILMIWMEDLVR,EDNFGACISLGK,LQDAYAGDK	AQSDSTAVAASR
FBpp0291042 Mhc	myosin heavy chain, isoform O	16	5	0	9.79	3.16	0	712	144	0.14	0.07	NDLENQLR,QLEEAESQVSQLSK,DIQTALEEEQR,NLADEVKDLLDQIGEGGR,ALDSMQASLEAEAK,ANALQNELEESR,LSIENSDLLR,LAGADIETYLLEK,IEELEEEVEAER,LSTEVEDLQLEVDR,LAEAEETIESLNQK,TALLDSLSGEK,KAMVDAAR,LTQEAVADLER,ELEELGER,LEQTLDELEDSLER	NDLENQLR,QLEEAESQVSQLSK,LTQEAVADLER,DIQTALEEEQR,NLADEVKDLLDQIGEGGR
FBpp0072127 Ca-P60A	calcium ATPase at 60A, isoform H	13	2	0	15.49	2.55	0	411.33	34.5	0.21	0.03	VIVITGDNK,EVFDSIVR,TGTLTTNQMSVSR,YGPNELPTEEGK,EFTLEFSR,LNSFSVNK,FSIPVVLLDETLK,SAAEMVLADDNFSSIVSAVEEGR,TVEQSLNFFGTDPER,EFDDLSPTEQK,VGEATETALIVLAEK,NILFSGTNVAAGK,LDEFGEQLSK	TGTLTTNQMSVSR,NILFSGTNVAAGK
FBpp0072339 Eps-15	epidermal growth factor receptor pathway substrate clone 15, isoform B	13	0	0	14.12	0	0	949		0.41		DPFGSDAFAANK,SMVTQLLETQR,GTGAIEAMTAAK,APSGEGFGDGGFVANFPK,AFDAFNDNFEDQFK,QSDLDKDGLVSGLEVK,SNTPLQNQR,YEDPFSIKDPFAEEGEEELSEGK,EIEELAR,SGLSDVVLSR,LDDLQAQVTK,AVVSVFDAFGEIGTR,AAMENQNAELVSEYQLK	
FBpp0312078 sesB	stress-sensitive B, isoform B	12	7	2	37.46	22.07	5.69	202.5	62	1.24	0.50	GTGGAFVLVLYDEIK,YFPTQALNFAFK,EFTGLGNCLTK,EQGFSSFWR,GMVDCFIR,GMLPDPK,TAVAPIER,QVFLGGVDK,ATEVIYK,GAFSNILR,QEGTGAFFK,SDGIVGLYR	YFPTQALNFAFK,EFTGLGNCLTK,EQGFSSFWR,TAVAPIER,GAFSNILR,QEGTGAFFK,SDGIVGLYR
FBpp0088481 Syn	synapsin, isoform F	9	0	0	12.54	0	0	469		0.32		TNQGSAMLEQITLTEK,DGFPLIEQTFFPNPR,DQGASILGAAVQSATQR,AGAGYQPVTNYEQQER,TPATAAPAPPR,NTDTLTER,TSPAVGSIGR,AGQRPPQTQNSVVEDAEDTMK,VQSLTGAAR	
FBpp0303494 CG31974	CG31974, isoform C	8	1	1	21.07	2.66	2.66	157.33	32	0.36	0.07	LVTSDER,DAVLVQENVTTR,YGEEGTPLK,FDMNSNIDQAETEIMNK,LVTSDERDVNR,DVDFSVLTK,VILADNSTIPK,LQLESGILPAQIFDGFPR	VILADNSTIPK
FBpp0085470 l(2)01289	lethal (2) 01289, isoform B	6	0	0	6.52	0	0	274		0.09		EWGIDEIPSIVLFER,TSDDDIDKEYDLPGLPALAFYR,ILNELENIDDELEKEGIVIVR,GIPTIYEGNLEDEEK,IPALYEGDLMNEDEVLEWLLVQK,STGDEDDVIEDVTSK	
FBpp0081031 Dap160	dynamin associated protein 160, isoform A	6	2	0	6.2	1.64	0	162	24	0.14	0.03	DTSMSEMSQLK,AELSALITK,KEDINTNDVQMSELK,ALQPQAGFVTGAQAK,LLQLTQER,YTQVFNANDR	SGYLTGSQAR,LLQLTQER
FBpp0293349 CG43078	CG43078, isoform B	5	2	1	3.16	1.17	0.7	80.5	89	0.07	0.02	VVMETIDDDEFFLR,QYISEAIR,GISEDNIQLR,QFAEFDEENR,AELDDNEDVGPR	QYISEAIR,AELDDNEDVGPR
FBpp0085357 Ars2	Ars2, isoform D	4	2	0	5.51	2.23	0	137	72	0.13	0.06	VAIADPLVER,FVQANTQELAK,VTNNDVQEYIK,ESIGGGGGDTLTQPAMLNLK	VDSSQADALIR,VAIADPLVER
FBpp0290458 CG9650	CG9650, isoform G	4	1	0	4.63	0.99	0	92	23	0.10	0.03	EAAAAAAAAAAAADNNNR,ENLASAMAAGMR,MEQLVSEQFR,TTSSGATSLVGELMDK	ENLASAMAAGMR
FBpp0305517 Atpalpha	Na pump alpha subunit, isoform K	3	0	0	3.89	0	0	45		0.05		FNTDDINFPIDNLR,DVSSDQLDEILR,EVSGDASEAALLK	
FBpp0307903 Ir25a	ionotropic receptor 25a, isoform E	1	1	0	0.64	0.64	0	51	52			QLEKLK	QLEKLK

gmr-gal4 > UAS-Flag-Strep-Nocte-Strep (*FSNS*), *gmr-gal4 > UAS-Flag-Strep-Nocte-HA* (*FSNH*) and control *gmr-gal4* (driver only, NEG). Data show the numbers of unique peptides, the % protein sequence coverage, Mascot Scores (Matrix Science) and EMPAI (empirical abundance index) scores and the peptides sequences derived from Mascot search engine. Data were compared using Protein Centre (Thermo). Black entries are high confidence hits and grey entries are lower confidence based on prior knowledge of known contaminants[14]. *nocte* mutants show defects in ChO morphology, pointing to a structural role of NOCTE in ChO cilia[9]. Consequently, the majority of the identified proteins (10/16) likely regulate function and dynamics of the ChO neuron cilia. As we were mainly interested in identifying potential temperature receptors, we focused on other NOCTE-interacting proteins, particularly on Ionotropic Receptor 25a (IR25a).

Extended Data Table 2 | Rhythm analysis of control and *IR25a* mutant flies under free running (DD) conditions at different ambient temperatures

Genotype	°C	n	%Rhythmic	Period(hr) ±SEM
Canton S	**18**	**32**	84	23.7± 0.2
***IR25a*$^{-/-}$**	**18**	**32**	75	24.1±0.3
rescue	**18**	**28**	60	23.6±0.1
***per*^{L}**	**18**	**22**	82	26.8±0.1
Canton-S	**25**	**45**	96	24.2±0.1
***IR25a*$^{-/-}$**	**25**	**44**	75	24.0±0.1
rescue	**25**	**30**	97	23.6±0.1
***per*^{L}**	**25**	**26**	38	28.4±0.2
Canton S	**29**	**60**	100	23.4±0.1
***IR25a*$^{-/-}$**	**29**	**39**	77	23.4±0.1
rescue	**29**	**27**	100	23.8±0.1
***per*^{L}**	**29**	**54**	52	29.6±0.4

Period values were calculated using autocorrelation as described in[40]. Flies with a rhythm statistics (RS) value >1.5 were considered rhythmic[40].

LETTER

doi:10.1038/nature15516

Fungal pathogen uses sex pheromone receptor for chemotropic sensing of host plant signals

David Turrà[1], Mennat El Ghalid[1], Federico Rossi[1] & Antonio Di Pietro[1]

For more than a century, fungal pathogens and symbionts have been known to orient hyphal growth towards chemical stimuli from the host plant[1,2]. However, the nature of the plant signals as well as the mechanisms underlying the chemotropic response have remained elusive[3]. Here we show that directed growth of the soil-inhabiting plant pathogen *Fusarium oxysporum* towards the roots of the host tomato (*Solanum lycopersicum*) is triggered by the catalytic activity of secreted class III peroxidases, a family of haem-containing enzymes present in all land plants[4]. The chemotropic response requires conserved elements of the fungal cell integrity mitogen-activated protein kinase (MAPK) cascade[5] and the seven-pass transmembrane protein Ste2, a functional homologue of the *Saccharomyces cerevisiae* sex pheromone α receptor[6]. We further show that directed hyphal growth of *F. oxysporum* towards nutrient sources such as sugars and amino acids is governed by a functionally distinct MAPK cascade. These results reveal a potentially conserved chemotropic mechanism in root-colonizing fungi, and suggest a new function for the fungal pheromone-sensing machinery in locating plant hosts in a complex environment such as the soil.

Root-colonizing fungi have a dramatic impact on plant health[7]. Beneficial symbionts such as mycorrhiza promote plant growth by supplying nutrients and microelements, while soil-borne pathogens provoke devastating yield losses and are highly persistent and difficult to control. *F. oxysporum* causes vascular wilt disease in over 100 field and greenhouse crops[8]. Infectious hyphae penetrate the roots preferentially through natural openings at the junctions of epidermal cells[9], indicating that the fungus can sense and grow towards chemical signals from the host plant. To learn more about the underlying mechanism, we developed a quantitative chemotropism assay on agar plates (Extended Data Fig. 1a–c). Microconidia of *F. oxysporum* exposed to a gradient of glutamate (Glu) produced significantly more germ tubes pointing towards the nutrient source than towards the solvent control, resulting in positive chemotropism (Extended Data Fig. 1d and Fig. 1a). Two hours exposure time was sufficient to induce a chemotropic response, indicating rapid reorientation of the hyphal growth axis towards the new gradient. Different nitrogen and carbon sources such as Glu, Asp or glucose elicited a chemotropic response, while others such as Gln, Met, ammonium or galactose did not (Fig. 1b). Importantly, germ tubes growing towards the chemoattractant did not differ in length from those growing towards the solvent, ruling out a bias from growth speed (Extended Data Fig. 1e). Thus, *F. oxysporum* responds rapidly and specifically to nutrients by redirecting hyphal growth towards the chemoattractant gradient.

In the model fungi *S. cerevisiae* and *Neurospora crassa*, chemotropism towards a mating partner is mediated by opposite gradients of diffusible peptide sex pheromones[6,10,11]. Although no sexual cycle has yet been reported in *F. oxysporum*, its genome encodes a putative protein with the characteristic hallmarks of fungal α-pheromone precursors, containing ten α-pheromone decapeptide repeats with near-identical sequence (Extended Data Fig. 1f, g). Synthetic α-pheromone from either

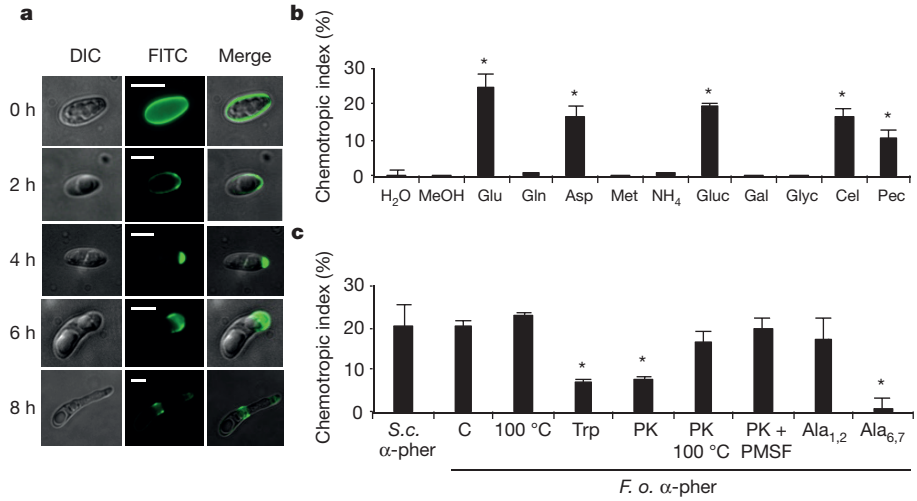


Figure 1 | *F. oxysporum* exhibits chemotropic growth towards different compounds. a, Germination of microconidia over time. Germ tube emergence sites are visualized by *Triticum vulgaris* lectin–fluorescein isothiocyanate (FITC) staining. DIC, differential interference contrast. Scale bar, 5 µm. **b**, Directed growth of germ tubes after 13 h exposure to a gradient of the indicated compounds. Gluc, glucose; Gal, galactose; Glyc, glycerol; Cel, cellulose; Pec, pectin (versus solvent control, $*P < 0.0001$). **c**, Directed growth towards a gradient of synthetic α-pheromone (α-pher) of *S. cerevisiae* (*S. c.*) or of *F. oxysporum* (*F. o.*), either untreated (C), boiled (100 °C) or treated with trypsin (Trp), proteinase K (PK), boiled PK (PK 100 °C) or PK plus its inhibitor phenylmethanesulfonylfluoride (PK + PMSF); or of α-pheromone analogues (D-Ala$_{1,2}$) or (D-Ala$_{6,7}$) (versus untreated, $*P < 0.0001$). **b**, **c**, Data are presented as the mean from two experiments. $n = 500$ germ tubes. Error bars show standard deviation (s.d.).

[1]Departamento de Genética, Campus de Excelencia Internacional Agroalimentario ceiA3, Universidad de Córdoba, 14071 Córdoba, Spain.

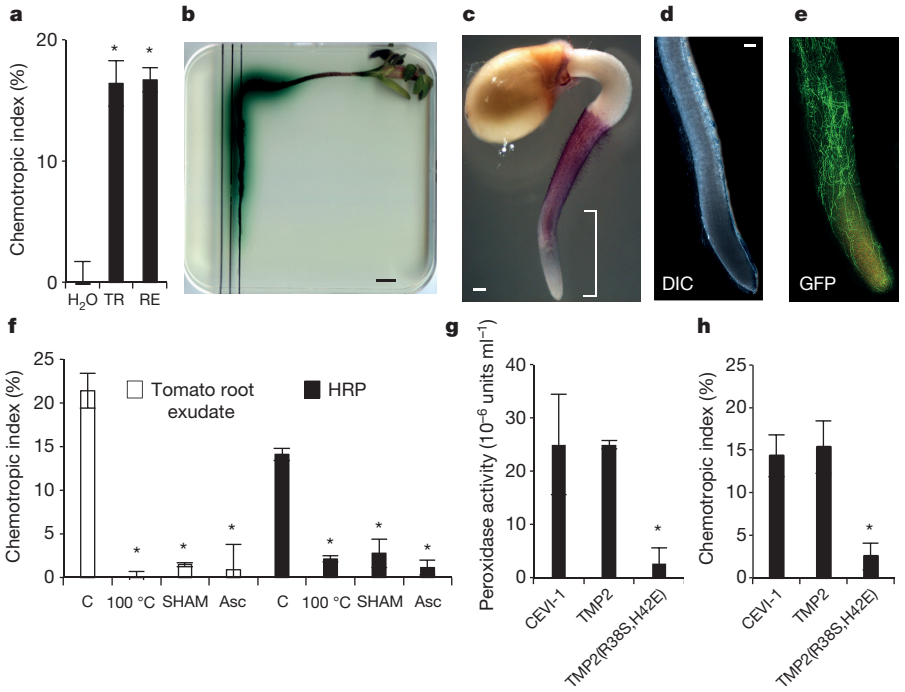


Figure 2 | Secreted tomato root peroxidases elicit fungal chemotropism. **a**, Directed growth of *F. oxysporum* towards tomato roots (TR) or root exudate (RE) (versus H_2O, $^*P < 0.0001$). **b**, **c**, Secretion of peroxidase activity (**b**) and its spatial distribution (**c**) in tomato roots was visualized by staining with 2,2′-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS) plus H_2O_2. Scale bars, 1 cm (**b**) and 1 mm (**c**). **d**, **e**, Detail of the root section marked in **c**, showing colonization by *F. oxysporum* expressing green fluorescent protein (GFP). Scale bar, 250 µm. Experiments were performed four times with similar results. **f**, Peroxidase enzymatic activity is required for chemoattraction. Directed growth of germ tubes towards root exudate or a gradient of 4 µM HRP, either untreated (C), boiled (100 °C), or in the presence of the peroxidase inhibitor salicylhydroxamic acid (SHAM) or the oxygen radical scavenger (+)-sodium L-ascorbate (Asc) (versus untreated, $^*P < 0.0001$). **g**, **h**, Chemoattractant activity of heterologously expressed tomato peroxidase requires an intact catalytic site. **g**, Enzymatic activity of 56 nM recombinant tomato peroxidases CEVI-1, TMP2 and TMP2(R38S,H42E), indicated as units ml^{-1}. Data are presented as the mean from two experiments, each with two technical replicates (versus TMP2, $^*P < 0.005$). **h**, Directed growth towards a gradient of 169 nM recombinant CEVI-1, TMP2 or TMP2(R38S,H42E) (versus TMP2, $^*P < 0.0001$). **a**, **f**, **h**, Data are presented as the mean from two experiments. $n = 500$ germ tubes. Error bars show s.d.

F. oxysporum or *S. cerevisiae* elicited a robust chemotropic response that was largely abolished by protease treatment or alanine substitution of two conserved residues (Gly_6 and Gln_7), indicating its specificity (Fig. 1c).

We next asked whether *F. oxysporum* exhibits directed growth towards the host plant. Both tomato roots and root exudate induced a significant chemotropic response (Fig. 2a). Chemoattractant activity of root exudate was sensitive to proteinase K treatment, partitioned into the water phase after ethyl acetate extraction and residing predominantly in the molecular weight fraction between 30 and 50 kilodaltons (kDa), suggesting that it originates from one or several proteins (Extended Data Fig. 2a). Separation by anion exchange chromatography (Extended Data Fig. 2b, c) and SDS–polyacrylamide gel electrophoresis (SDS–PAGE) identified two protein bands in the chemotropically active fractions that were absent from the inactive ones and elicited a significant chemotropic response (Extended Data Fig. 2d, e). Analysis by in-gel tryptic digestion followed by liquid chromatography-electrospray ionization-tandem mass spectrometry (LC-ESI-MC/MC) identified three tomato proteins, TMP1, TMP2 and CEVI-1 (Extended Data Fig. 3a). They belong to class III peroxidases, secreted haem-containing oxidoreductases present in all land plants, which catalyse the reductive cleavage of hydrogen peroxide by an electron donor[12], are encoded by multigene families, and function in diverse physiological processes such as cell wall modification and pathogen defence[4]. *Arabidopsis thaliana* has over 70 members of this family, many of which are expressed and secreted in roots[13].

We observed a strong gradient of peroxidase activity exuded by tomato roots into the adjacent medium (Fig. 2b). The highest enzymatic activity was associated with the root hair zone, which also exhibits the highest density of colonization by *F. oxysporum* (Fig. 2c–e). Secreted peroxidase activity differed considerably between root exudates collected from different tomato plants, reflecting the multiplicity and complex regulation of plant class III peroxidase genes[4]. Importantly, enzymatic activity in individual root exudates correlated significantly with fungal chemoattraction (Extended Data Fig. 2f, g). We next tested horseradish peroxidase (HRP), which shares 39%, 37% and 49% amino acid identity with tomato TMP1, TMP2 and CEVI-1, respectively (Extended Data Fig. 3b), and whose molecular structure and catalytic properties are well characterized[12]. Commercial HRP triggered a robust chemotropic response in *F. oxysporum* (Fig. 2f). Chemotropism induced by HRP and tomato root exudate was abolished when peroxidase enzymatic activity was eliminated by boiling or by addition of the specific inhibitor salicylhydroxamic acid (SHAM), or in the presence of the oxygen radical scavenger ascorbate (Extended Data Fig. 2h and Fig. 2f). However, these inhibitors did not prevent chemotropism towards glucose or α-pheromone (Extended Data Fig. 2i).

To confirm the importance of peroxidase catalytic activity in the process of chemoattraction, we heterologously expressed tomato peroxidases CEVI-1 and TMP2, as well as a point-mutated version of TMP2 in which the conserved Arg 38 and His 42 residues[12] were substituted by Ser and Glu, respectively (Extended Data Fig. 3b). Recombinant catalytically active CEVI-1 and TMP2, but not catalytically inactive TMP2(R38S,H42E) triggered a robust chemotropic response in *F. oxysporum* (Fig. 2g, h). Collectively, these findings demonstrate that secreted root peroxidases elicit a chemotropic response in *F. oxysporum* through a mechanism that requires enzymatic activity.

We next asked whether directed hyphal growth towards different chemoattractants is mediated by common or distinct cellular mechanisms. Both α-pheromone and Glu exhibited a bell-shaped dose–response curve with a gradual decrease at higher concentrations (Fig. 3a), which could be explained by receptor saturation, as previously shown

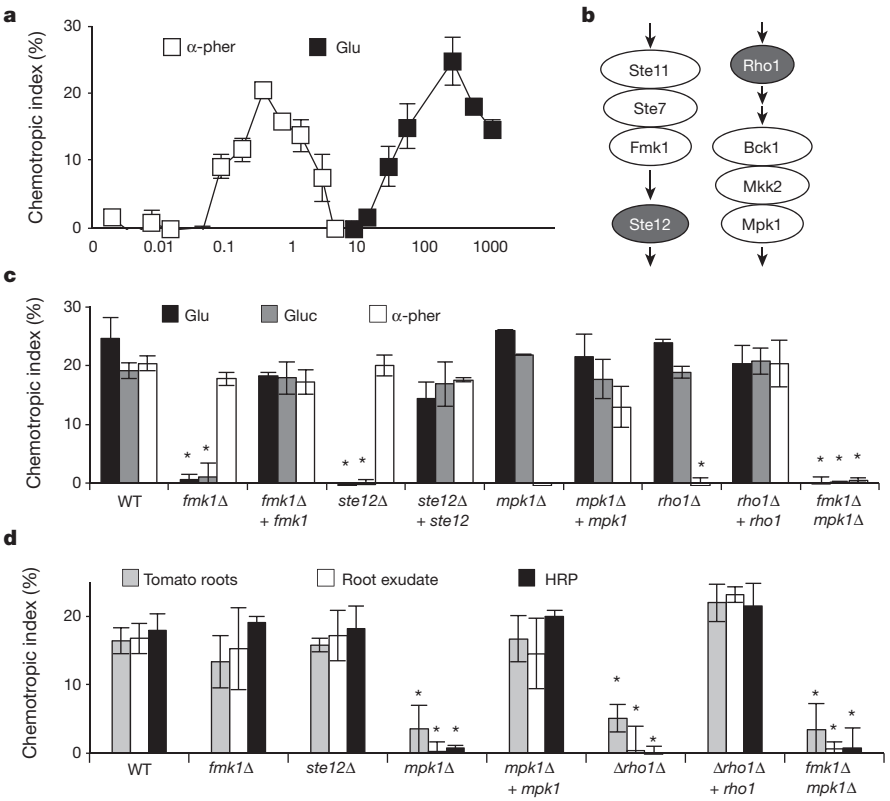


Figure 3 | Chemotropism towards nutrients and α-pheromone is governed by distinct MAPK cascades. a, Dose–response curves for chemotropism towards α-pheromone (α-pher) or Glu. **b**, Elements of the *F. oxysporum* Fmk1 and Mpk1 MAPK cascades. **c**, **d**, Directed growth of indicated fungal strains towards a gradient of Glu, glucose (Gluc) or α-pheromone (**c**), or tomato roots, root exudate or HRP (**d**) (versus wild type (WT), $*P < 0.0001$). **a**, **c**, **d**, Data are presented as the mean from two experiments. $n = 500$ germ tubes. Error bars show s.d.

for the pheromone response in *S. cerevisiae*[11]. Remarkably, chemotropic sensitivity of *F. oxysporum* to α-pheromone was three orders of magnitude higher compared to Glu, suggesting that the two responses may be governed by distinct cellular mechanisms. To test this idea, we used fungal mutants lacking defined elements of MAPK cascades, three-component signalling modules conserved from yeast to humans that function in succession to transmit a variety of cellular signals[14]. Like most fungi, *F. oxysporum* has three MAPKs orthologous to *S. cerevisiae* Kss1, Mpk1 and Hog1 (ref. 15). The p42/44 MAPK Fmk1, its upstream MAPKK Ste7 and MAPKKK Ste11, as well as the downstream transcription factor Ste12 (Fig. 3b), function in a conserved pathway that governs filamentation in *S. cerevisiae* and invasive growth in plant pathogens[15–17]. Isogenic *F. oxysporum* mutants lacking *fmk1*, *ste7*, *ste11* or *ste12* were impaired in chemotropism towards Glu or glucose, but not α-pheromone (Fig. 3c and Extended Data Fig. 4). *S. cerevisiae* mutants lacking the orthologous MAPKs Fus3 or Kss1 also maintain most of the chemotropic response towards α-pheromone[6]. To investigate whether another MAPK mediates chemotropic sensing of α-pheromone, we used the chemical inhibitors PD98059 and SB202190, which selectively block p42/44 MAPKs and p38 MAPKs, respectively. PD98059, but not SB202190, prevented chemotropism towards α-pheromone (Extended Data Fig. 5a), pointing to a role of the second p42/44 MAPK, Mpk1. In *S. cerevisiae,* Mpk1 together with Mkk1/2 and Bck1 functions in the cell wall integrity (CWI) MAPK module[5], which is activated during the pheromone response[18]. Mutations in CWI pathway components have pleiotropic effects on cell wall architecture and impair fungal virulence on plants[5,15,19]. Loss of the orthologous genes *mpk1*, *mkk2* or *bck1* in *F. oxysporum* (Fig. 3b) led to high sensitivity against the cell-wall-perturbing compounds Congo red and Calcofluor white, confirming that they function in the CWI response (Extended Data Fig. 5b–g). Interestingly, the mutants were impaired in chemotropism towards α-pheromone but remained responsive to Glu or glucose (Fig. 3c and Extended Data Fig. 5h, i). Moreover, a mutant in the small G protein Rho1 (ref. 19), located upstream of the CWI MAPK module[5], also failed to grow towards α-pheromone. In *S. cerevisiae,* Rho1 mediates localization of key components of the CWI pathway to the tips of pheromone-induced mating projections[20].

We sought to confirm these results independently using a different chemotropism assay based on the angle of hyphal tip projections relative to the chemoattractant gradient. This method was used extensively in *S. cerevisiae* to study chemotropic responses to mating pheromones[6,10,11]. The average cosine of hyphal tip projection angles was significantly higher when *F. oxysporum* was exposed to a gradient of Glu or α-pheromone compared with the water control, confirming positive chemotropism (Extended Data Fig. 6a, b). The *fmk1*Δ mutant was specifically impaired in growth towards Glu while the *mpk1*Δ mutant failed to respond to α-pheromone. Taken together, these results establish that the invasive growth and CWI MAPK pathways have distinct and complementary roles in chemotropic sensing of nutrients and sex pheromones. Consistent with this model, a *fmk1*Δ *mpk1*Δ double mutant responded neither to nutrients nor to α-pheromone (Fig. 3c).

We next tested whether these MAPK cascades are required for chemotropic growth of *F. oxysporum* towards the host plant. Mutants lacking Fmk1, Ste7, Ste11 or Ste12 were not affected in chemotropism towards tomato roots or root exudate, but those lacking Mpk1, Mkk2, Bck1 or Rho1 were impaired (Fig. 3d and Extended Data Figs 4g, 5i). Importantly, CWI components were also essential for the chemotropic response to HRP while Fmk1 and Ste12 were not (Fig. 3d and Extended Data Fig. 5i). Thus, peroxidase-mediated chemotropism of *F. oxysporum* towards plant roots is specifically governed by the CWI MAPK cascade.

In *S. cerevisiae* and *N. crassa*, chemotropic sensing of α-pheromone requires the seven-pass transmembrane (7TM) G-protein-coupled receptor (GPCR) Ste2 or Pre-2, respectively[6,10]. *F. oxysporum* has a putative Ste2 orthologue with seven predicted transmembrane regions and a topology characteristic of ascomycete α-pheromone receptors (Extended Data Fig. 7a, b). Loss of Ste2 (Extended Data Fig. 7c, d)

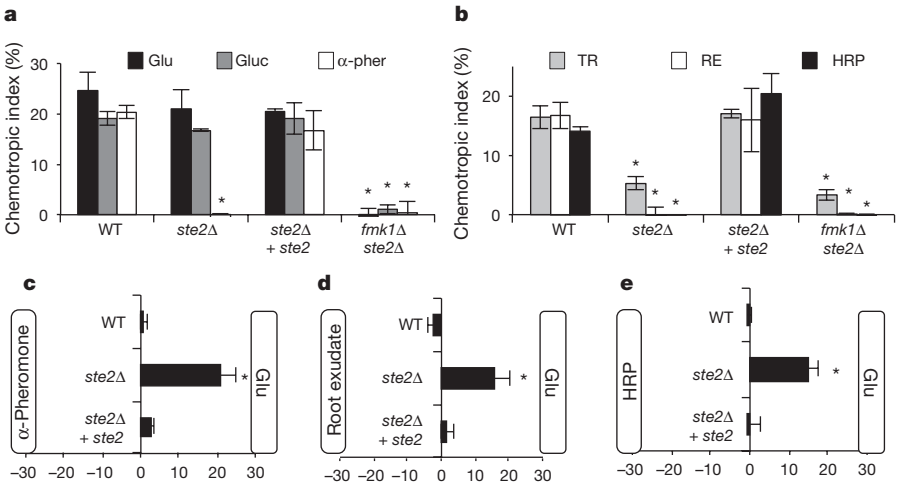


Figure 4 | The 7TM domain receptor Ste2 is required for chemotropic sensing of α-pheromone and root peroxidases. a–e, Directed growth of the indicated fungal strains exposed to a gradient of Glu, glucose (Gluc) or α-pheromone (α-pher) (**a**); tomato roots (TR), root exudate (RE) or HRP (**b**); or exposed simultaneously to competing gradients of Glu opposed to either α-pheromone (**c**), root exudate (**d**) or HRP (**e**) (versus wild type, $*P < 0.0001$). Data are presented as the mean from two experiments. $n = 500$ germ tubes. Error bars show s.d.

abolished chemotropism of *F. oxysporum* towards α-pheromone but not towards nutrients, confirming that it is a functional homologue of yeast Ste2 (Fig. 4a). Strikingly, *ste2Δ* mutants were impaired in chemotropism towards tomato roots, root exudate and HRP (Fig. 4b). This was unexpected because Ste2 is generally regarded as a specific receptor for α-pheromone[21]. To corroborate further the role of Ste2 in chemotropic sensing of root chemoattractants, fungal strains were exposed to opposite gradients of Glu versus either α-pheromone, root exudate or HRP. The wild-type and the complemented strain failed to display directed growth, suggesting that chemotropism is annulled in the presence of two competing gradients (Fig. 4c–e), while *ste2Δ* grew towards Glu, confirming its incapacity to sense a competing gradient of α-pheromone, root exudate or HRP. Loss of Ste2 caused a small but significant decrease in virulence of *F. oxysporum* on tomato plants, indicating that Ste2-mediated chemotropic sensing of secreted root compounds is important for initiation of fungal infection (Extended Data Fig. 7e). Interestingly, in a previous study *Petunia* plants defective in secretion of the signalling compound strigolactone were significantly delayed in the symbiotic interaction with arbuscular mycorrhizal fungi[22].

Our results reveal a previously unknown ability of the fungal pathogen *F. oxysporum* to reorient hyphal growth towards a variety of chemical signals. On the basis of genetic evidence, we propose that chemotropism is mediated by distinct MAPK modules: Fmk1 for nutrients and Mpk1 for sex pheromones and plant compounds (Extended Data Fig. 8). Remarkably, *F. oxysporum* uses the same signalling pathway for chemotropic sensing of mating factors and host cues, including Ste2, a 7TM GPCR that was previously thought to function specifically in α-pheromone sensing. How secreted plant peroxidases generate a chemoattractant signal, and how Ste2 mediates signal sensing and chemotropic response in concert with the CWI pathway, remain to be determined. Since class III peroxidases and fungal MAPK cascades are evolutionarily conserved[4,15], our findings might be of general relevance to the chemotropic interaction between plants and root-colonizing fungi.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 17 June 2014; accepted 21 August 2015.
Published online 26 October 2015.

1. de Bary, A. *Vergleichende Morphologie und Biologie der Pilze, Mycetozoen, und Bacterien* (Wilhelm Engelmann, 1884).
2. Zentmyer, G. A. Chemotaxis of zoospores for root exudates. *Science* **133,** 1595–1596 (1961).
3. Brand, A. & Gow, N. A. Mechanisms of hypha orientation of fungi. *Curr. Opin. Microbiol.* **12,** 350–357 (2009).
4. Passardi, F., Penel, C. & Dunand, C. Performing the paradoxical: how plant peroxidases modify the cell wall. *Trends Plant Sci.* **9,** 534–540 (2004).
5. Levin, D. E. Cell wall integrity signaling in *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.* **69,** 262–291 (2005).
6. Arkowitz, R. A. Chemical gradients and chemotropism in yeast. *Cold Spring Harb. Perspect. Biol.* **1,** a001958 (2009).
7. Berendsen, R. L., Pieterse, C. M. & Bakker, P. A. The rhizosphere microbiome and plant health. *Trends Plant Sci.* **17,** 478–486 (2012).
8. Dean, R. *et al.* The Top 10 fungal pathogens in molecular plant pathology. *Mol. Plant Pathol.* **13,** 414–430 (2012).
9. Pérez-Nadales, E. & Di Pietro, A. The membrane mucin Msb2 regulates invasive growth and plant infection in *Fusarium oxysporum*. *Plant Cell* **23,** 1171–1185 (2011).
10. Kim, H. & Borkovich, K. A. A pheromone receptor gene, *pre-1*, is essential for mating type-specific directional growth and fusion of trichogynes and female fertility in *Neurospora crassa*. *Mol. Microbiol.* **52,** 1781–1798 (2004).
11. Segall, J. E. Polarization of yeast cells in spatial gradients of alpha mating factor. *Proc. Natl Acad. Sci. USA* **90,** 8332–8336 (1993).
12. Berglund, G. I. *et al.* The catalytic pathway of horseradish peroxidase at high resolution. *Nature* **417,** 463–468 (2002).
13. Badri, D. V. *et al.* Root secreted metabolites and proteins are involved in the early events of plant-plant recognition prior to competition. *PLoS ONE* **7,** e46640 (2012).
14. Widmann, C., Gibson, S., Jarpe, M. B. & Johnson, G. L. Mitogen-activated protein kinase: conservation of a three-kinase module from yeast to human. *Physiol. Rev.* **79,** 143–180 (1999).
15. Turrà, D., Segorbe, D. & Di Pietro, A. Protein kinases in plant-pathogenic fungi: conserved regulators of infection. *Annu. Rev. Phytopathol.* **52,** 267–288 (2014).
16. Di Pietro, A., García-MacEira, F. I., Méglecz, E. & Roncero, M. I. A MAP kinase of the vascular wilt fungus *Fusarium oxysporum* is essential for root penetration and pathogenesis. *Mol. Microbiol.* **39,** 1140–1152 (2001).
17. Liu, H., Styles, C. A. & Fink, G. R. Elements of the yeast pheromone response pathway required for filamentous growth of diploids. *Science* **262,** 1741–1744 (1993).
18. Buehrer, B. M. & Errede, B. Coordination of the mating and cell integrity mitogen-activated protein kinase pathways in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **17,** 6517–6525 (1997).
19. Martínez-Rocha, A. L. *et al.* Rho1 has distinct functions in morphogenesis, cell wall biosynthesis and virulence of *Fusarium oxysporum*. *Cell. Microbiol.* **10,** 1339–1351 (2008).
20. Bar, E. E., Ellicott, A. T. & Stone, D. E. Gβγ recruits Rho1 to the site of polarized growth during mating in budding yeast. *J. Biol. Chem.* **278,** 21798–21804 (2003).
21. Xue, C., Hsueh, Y. P. & Heitman, J. Magnificent seven: roles of G protein-coupled receptors in extracellular sensing in fungi. *FEMS Microbiol. Rev.* **32,** 1010–1032 (2008).
22. Kretzschmar, T. *et al.* A petunia ABC protein controls strigolactone-dependent symbiotic signalling and branching. *Nature* **483,** 341–344 (2012).

Acknowledgements The authors are grateful to E. Martínez Aguilera for technical assistance. This work was supported by grants BIO2010-15505 and BIO2013-47870-R from the Spanish Ministerio de Innovación y Competitividad (MINECO), and BIO2008-04479 from MINECO/ERA-NET PathoGenoMics to A.D.P. M.E.G. was supported by the Marie Curie ITN ARIADNE (FP7-PEOPLE-ITN-237936) from the European Commission. F.R. had a fellowship from the ERASMUS student exchange program.

Author Contributions D.T. and A.D.P. initiated the work and designed the experiments. D.T., M.E.G. and F.R. carried out the experiments and analysed the data. D.T. and A.D.P. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.D.P. (ge2dipia@uco.es).

METHODS

Fungal strain culture and transformation. Fungal strains used in this study are listed in Extended Data Table 1. All are derivatives of *F. oxysporum* f. sp. *lycopersici* isolate 4287 (FGSC 9935). Strain culture and storage were performed as described[16]. Phenotypic analysis of colony growth and invasion of cellophane membranes was done as reported[23]. Targeted gene replacement with the hygromycin resistance cassette and complementation of the mutants by co-transformation with the phleomycin resistance cassette were performed as reported[23]. Oligonucleotides used to generate PCR fragments for gene replacement, mutant identification and complementation are listed in Extended Data Table 2. *F. oxysporum* gene data are available in the *Fusarium* Comparative Database at the Broad Institute under the following accession numbers: *ste11*, FOXG_09411; *ste7*, FOXG_05521; *fmk1*, FOXG_08140; *ste12*, FOXG_02103; *ste2*, FOXG_10633; *rho1*, FOXG_13835; *bck1*, FOXG_08078; *mkk2*, FOXG_02117; *mpk1*, FOXG_05092.

Quantification of fungal chemotropism. Freshly obtained microconidia were embedded in 4 ml water agar (WA; 0.5%, w/v) (Oxoid) at a final concentration of 2.5×10^6 per ml and poured into a standard Petri dish (Extended Data Fig. 2a). A central scoring line was drawn on the bottom of the plate, and two parallel wells were cut into the WA layer on both sides at 5 mm distance from the scoring line. Then, 50 µl of the test compound solution or the solvent control were added to the wells at both sides of the scoring line. In gradient competition experiments, solutions of the two different test compounds were applied at both sides of the scoring line. Tested compounds and standard concentrations were: sodium glutamate (Glu), glutamine (Gln), sodium aspartate (Asp), methionine (Met), all at 295 mM; ammonium nitrate (NH_4,), glucose (Gluc), galactose (Gal), glycerol (Glyc), all at 50 mM; or cellulose (Cel), pectin (Pec), all at 1% (w/v). Sterile water or methanol were used as solvent controls. To measure chemotropism towards tomato plants, the root of a 2-week-old tomato seedling was placed directly on top of one of the wells. A sterile metal string was placed on the opposite well as a control. Plates were maintained in a plastic box at 28 °C in the dark for the indicated time periods (13 h unless otherwise stated). Chemotropism of conidial germ tubes was quantified with an Olympus binocular microscope (200× magnification), by counting the number of hyphal tips pointing towards the test compound and those pointing towards the solvent control. The chemotropic index was calculated as $((H_{test} - H_{solv})/H_{total} \times 100)$, where H_{test} is the number of hyphae growing towards the test compound, H_{solv} is the number of hyphae growing towards the solvent control, and H_{total} is the total number of hyphae counted. For each test compound a total of 500 hyphal tips were scored. All experiments were performed at least twice. Statistical analysis was conducted using *t*-test.

For the hyphal tip projection assay, light microscopy photographs of chemotropism plate assays were recorded in a Leica DMR microscope (200× magnification) using a Leica DFC 300 FX digital camera, and the angle (in degrees) of the hyphal tip relative to the chemoattractant gradient was measured using the ImageJ software[24]. For each test compound a total of 300 hyphal tip projection angles were measured. All experiments were performed at least twice. Length of germ tubes growing towards the test compound or the solvent control was measured using ImageJ. For visual monitoring of compound diffusion, 50 µl of a 1% (w/v) solution of Congo red in water was added to the test compound well and 50 µl of water into the solvent control well. Plates were incubated at 28 °C, and dye diffusion was documented after different time periods in a Leica MZ FLIII fluorescence stereomicroscope using a Leica DFC 300 FX digital camera. Dye intensity was quantified with the KODAK 1D Image Analysis software.

The MAPK inhibitors PD98059 and SB202190 (Calbiochem) were added to the WA medium at a final concentration of 10 or 30 µM, respectively, before adding the chemoattractant compound. Commercial horseradish peroxidase (HRP; Sigma) was assayed at a standard concentration of 4 µM. Peroxidase inhibitors/scavengers salicylhydroxamic acid (SH), thiourea (TU) and (+)-sodium L-ascorbate (Asc) (Sigma) were added directly to the chemoattractant solution at a final concentration of 60 mM, 60 mM and 160 mM, respectively. Synthetic *F. oxysporum* α-pheromone, its analogues (D-$Ala_{1,2}$) and (D-$Ala_{6,7}$), and *S. cerevisiae* α-factor were obtained from GenScript (Piscataway). Lyophilized peptides were dissolved in 50% (v/v) methanol in water and assayed at a standard concentration of 378 µM. To test the effect of different treatments on the chemoattractant activity of *F. oxysporum* α-pheromone, the peptide was incubated for 10 min at 100 °C, or for 30 min at 37 °C with 1 mg ml^{-1} trypsin, 1 mg ml^{-1} proteinase K, 1 mg ml^{-1} heat-denatured (10 min at 100 °C) proteinase K, or 1 mg ml^{-1} proteinase K plus 1 mM phenylmethylsulfonyl fluoride (PMSF) (all from Sigma).

Light and fluorescence microscopy. Low-resolution imaging was performed using a Lumar.V12 fluorescence stereomicroscope (Zeiss). Wide-field fluorescence imaging was performed using a Zeiss Axio Imager M2 microscope equipped with a Photometrics Evolve EMCCD camera. To visualize growth of *F. oxysporum* on tomato roots, freshly obtained microconidia of the wild-type–GFP strain[16] were embedded in WA as described earlier. One-day-old germinated tomato seeds were placed on top of the medium and incubated for 2 days at 28 °C before observation in the microscope.

To visualize sites of germ tube emergence, *F. oxysporum* microconidia were stained with 50 µg ml^{-1} fluorescein isothiocyanate-labelled lectin from *Triticum vulgaris* (WGA-FITC) (Sigma) in PBS containing 12.5 µM $CaCl_2$ and 12.5 µM $MnCl_2$.

Identification of the root chemoattractant. Tomato seeds (cultivar Monica) were provided as a gift by Syngenta Seeds. Seeds were surface sterilized[23], planted in moist vermiculite and maintained in a growth chamber (15/9 h light/dark photoperiod, 28 °C) until the plants reached the second true leaf stage. Roots were washed carefully to remove the adhering substrate, placed in sterile water and kept at 25 °C for 48 h. The collected root exudate was filtered through a 0.22-µm Millipore membrane and stored at −20 °C until use. To measure fresh root weight, roots were cut from individual plants, gently blotted with a paper towel and weighed.

Filter-sterilized root exudates were partitioned with ethyl acetate to obtain an ethyl acetate fraction and a water fraction (WF). The WF was further separated by centrifugal ultrafiltration (MWCO 10 kDa; 30 kDa; 50 kDa) (Corning). The 30–50 kDa fraction was applied to a Hitrap QFF anion exchange chromatography column on an ÄKTA purifier (GE Healthcare), and proteins were eluted with a linear gradient of NaCl. Fractions were desalted by dialysis, tested for chemotropic activity as described earlier and analysed by SDS–polyacrylamide (10%) gel electrophoresis, followed by Coomassie blue staining. Proteins bands were eluted from the gel and tested for chemoattractant activity. Proteins of interest were subjected to tryptic digestion and analysed by liquid chromatography-electrospray ionization-mass spectrometry (LC-ESI-MS). Identification of tomato proteins was carried out at the Protein Micro-Analysis Core Facility of the Biozentrum, Innsbruck Medical University.

Heterologous expression of tomato peroxidases. Recombinant tomato peroxidase proteins were produced in *Escherichia coli* strain BL21 (DE3) ung-151 transformed with plasmids pET28a-Cevi-1, pET28a-Tmp2 and pET28a-Tmp2(R38A,H42A), respectively. Plasmids were obtained by subcloning the corresponding complementary DNA fragments lacking the sequence encoding the signal peptide in the vector pET28a(+), using XhoI and NdeI. Solubilization and re-folding of recombinant peroxidases was performed as described[25]. Purification of the recombinant proteins was performed on an ÄKTA purifier using a Ni-NTA chromatography column.

Peroxidase enzymatic activity assays. To visualize peroxidase activity secreted by roots, 4-day-old tomato seedlings were placed on 0.5% WA supplemented with 0.91 mM 2,2′-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid (ABTS) (Sigma) and 2.5 mM H_2O_2 (J. T. Baker) (WA-ABTS), and incubated for 45 min at 28 °C. Seedlings were carefully removed from WA-ABTS, placed in a Petri dish containing 0.5% agarose (w/v) and imaged with a stereo microscope.

Peroxidase activity assays were carried out in 96-well microtitre plates. The reaction mixture contained 0.91 mM ABTS, 2.5 mM H_2O_2 in phosphate-citrate buffer (51 mM Na_2HPO_4. 24 mM citric acid, pH 5.6) in a final volume of 150 µl. Where applicable, peroxidase inhibitors/scavengers (75 µM TU and SH or 250 µM Asc) were pre-incubated with the buffer for 5 min before adding ABTS and H_2O_2. For each reaction, a blank containing heat-inactivated (20 min boiling) peroxidase was included. Reactions were incubated at 28 °C and absorbance at 405 nm was measured at different time intervals in a Spectrafluor Plus microplate reader (Tecan). Peroxidase activity was calculated in units per ml, using the formula $((\Delta A_{405\,nm}/\text{min test} - \Delta A_{405\,nm}/\text{min blank}) \times (\text{total volume assay}) \times (\text{dilution factor}))/((\text{millimolar extinction coefficient of oxidized ABTS at 405 nm}) \times (\text{volume enzyme used}))$. Statistical analysis was conducted using *t*-tests.

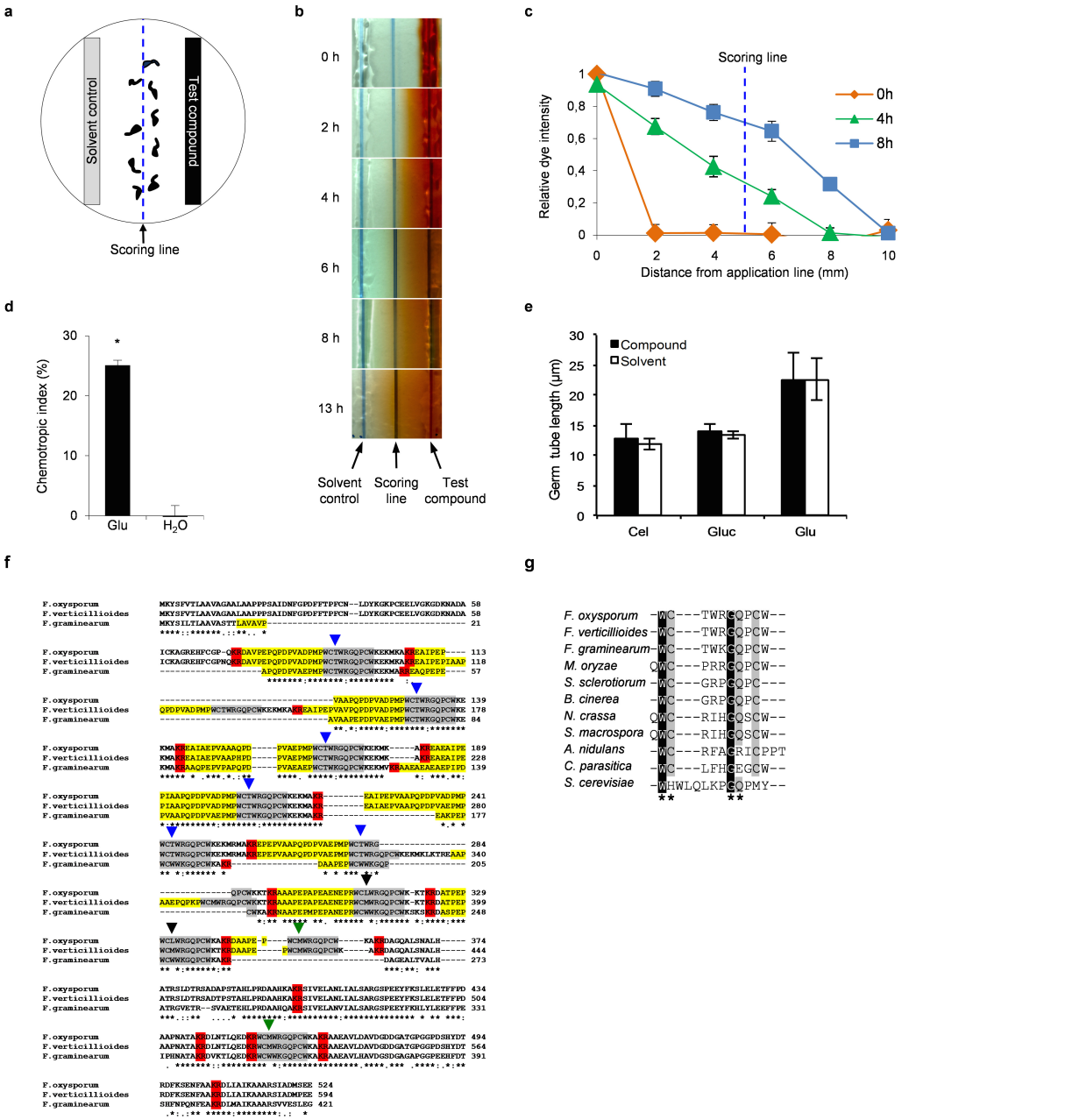
Tomato seedling infection assay. Surface-sterilized tomato seeds (cultivar Monica) were transferred to sterile glass tubes containing WA or WA supplemented with 2.5×10^6 microconidia per ml of the different *F. oxysporum* strains. Plants were maintained in a growth chamber (15/9 h light/dark cycle, 28 °C). Survival was recorded daily, calculated by the Kaplan–Meier method and compared among groups using the log-rank test. Virulence experiments were conducted with 40 plants per treatment and performed twice with similar results.

Bioinformatic and statistical analysis. *F. oxysporum* predicted proteins Ste11, Ste7, Bck1, Mkk2, Mpk1 and Ste2 were identified by BLASTp search of the *Fusarium* Comparative Database at the Broad Institute (http://www.broadinstitute.org/annotation/genome/fusarium_group/MultiHome.html), using the amino acid sequences of the *S. cerevisiae* proteins. Protein alignments and phylogenetic comparisons were done using ClustalW (ref. 26) and MEGA5 (ref. 27). Protein domain predictions were made using the Prosite database (ExPASy; Swiss

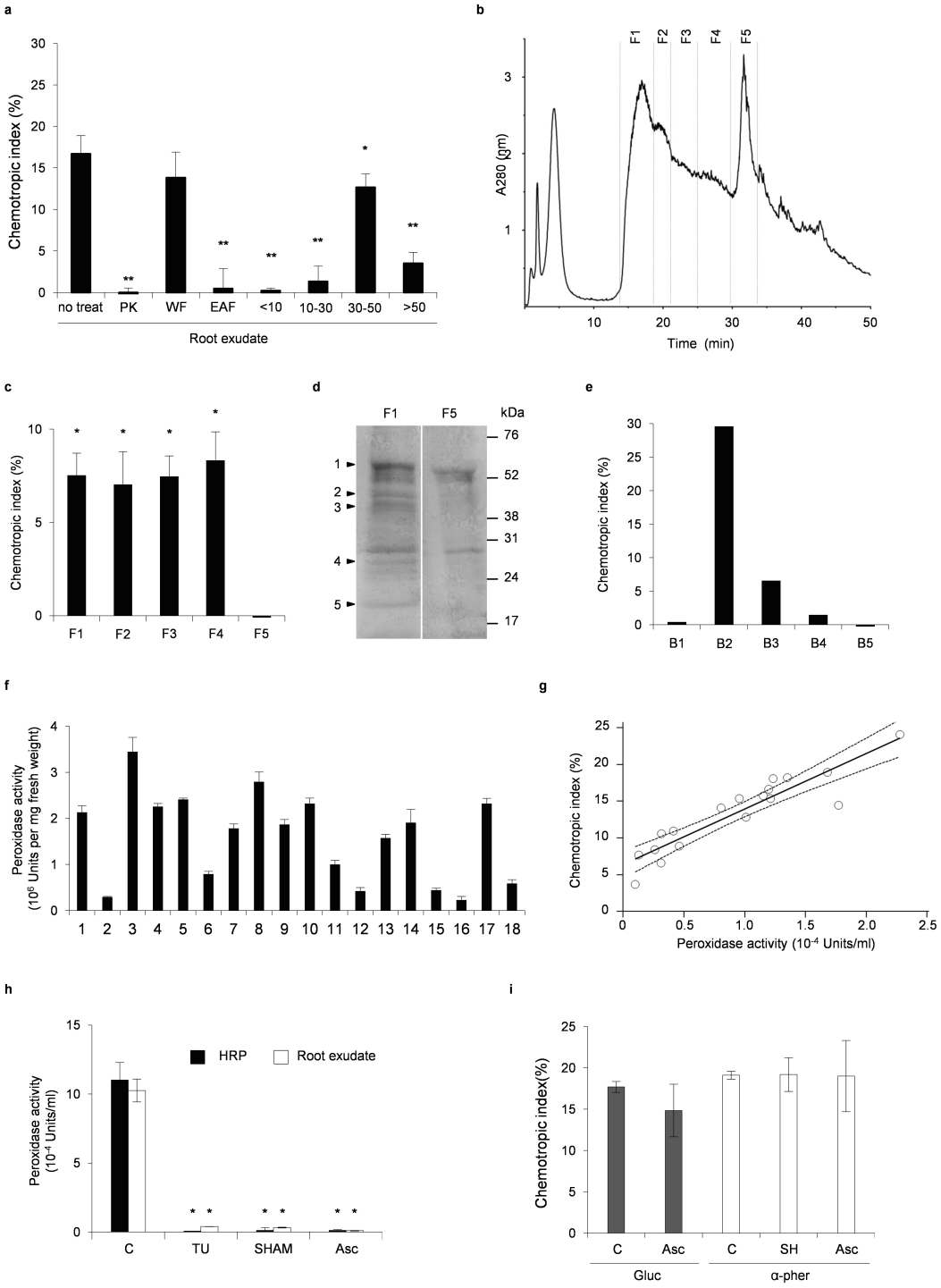
Institute of Bioinformatics). Prediction of Ste2 transmembrane helices was done with SOSUI[28]. Linear regression analysis was conducted using MedCalc v. 12.1.0 (MedCalc Software).

Data reporting. No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

23. López-Berges, M. S., Rispail, N., Prados-Rosales, R. C. & Di Pietro, A. A nitrogen response pathway regulates virulence functions in *Fusarium oxysporum* via the protein kinase TOR and the bZIP protein MeaB. *Plant Cell* **22,** 2459–2475 (2010).
24. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671–675 (2012).
25. Smith, A. T., Sanders, S. A., Thorneley, R. N., Burke, J. F. & Bray, R. R. Characterisation of a haem active-site mutant of horseradish peroxidase, Phe41→Val, with altered reactivity towards hydrogen peroxide and reducing substrates. *Eur. J. Biochem.* **207,** 507–519 (1992).
26. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23,** 2947–2948 (2007).
27. Tamura, K. *et al.* MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28,** 2731–2739 (2011).
28. Hirokawa, T., Boon-Chieng, S. & Mitaku, S. SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics* **14,** 378–379 (1998).
29. Rispail, N., & Di Pietro, A. *Fusarium oxysporum* Ste12 controls invasive growth and virulence downstream of the Fmk1 MAPK cascade. *Mol Plant Microbe Interact* **22,** 830–839 (2009).



Extended Data Figure 1 | Plate assay for quantitative determination of directed hyphal growth and identification of a *F. oxysporum* orthologue of the *S. cerevisiae* α-pheromone precursor. a, Schematic representation of the plate chemotropism assay. Test compound and solvent control are applied to opposite sides of a Petri dish containing a layer of water agar with $2.5 \times 10^6\,ml^{-1}$ *F. oxysporum* microconidia, at a distance of 0.5 cm from the central scoring line. Chemotropic index was calculated as $((H_{test} - H_{solv})/H_{total} \times 100)$, where H_{test} is the number of hyphae growing towards the test compound, H_{solv} is the number of hyphae growing towards the solvent control, and H_{total} is the total number of hyphae counted. **b**, Visualization of compound diffusion and gradient establishment. The dye Congo red (1% w/v in water) was loaded into the application well on the right side of the scoring line. Diffusion was recorded photographically after the indicated time intervals. **c**, Dye intensity in experiment **b** was measured at the indicated distances from the application well after different time intervals, using the Kodak Image Analyzer software. The blue dashed line represents the relative position of the scoring line. Mean values were calculated from measurements of five individual spots per distance. **d**, Direction of germ tube emergence after 2 h exposure to a gradient of Glu or the solvent (H_2O) was quantitatively determined by lectin–FITC staining and expressed as chemotropic index (versus H_2O, $^*P < 0.0001$). Data are presented as the mean from two experiments. $n = 200$ germ tube emergence sites. **e**, Lengths of germ tubes exposed for 13 h to a gradient of 1% (w/v) cellulose (Cel), 55 mM glucose (Gluc), 295 mM Glu or the solvent (H_2O) were measured using the ImageJ software. The mean length of germ tubes growing towards the nutrient chemoattractants is not significantly different from that of germ tubes growing towards the solvent. Data are presented as the mean from two experiments. $n = 100$ germ tubes. Error bars show s.d. **f**. The predicted product of the *F. oxysporum* α-pheromone precursor gene (*Fusarium* Comparative Database accession FOXG_08636) was aligned with predicted α-pheromone precursors from *F. graminearum* (FGSG_05061) and *F. verticillioides* (FVEG_06038). Conserved residues are indicated with an asterisk. Predicted KR and RR cleavage signals for KEX2-like endopeptidases are highlighted in red. Predicted maturation signals characterized by the presence of XA or XP dipeptide repeats are highlighted in yellow. Predicted mature α-pheromone decapeptide repeats are highlighted in grey. Coloured arrowheads indicate differences between the decapeptide repeats at the third amino acid residue. **g**, Amino acid alignment of predicted mature α-pheromone of *F. oxysporum* with orthologues from ascomycete fungi. Absolutely and highly conserved residues are shaded in black and in grey, respectively. Residues replaced with alanines in the ($Ala_{1,2}$) or ($Ala_{6,7}$) analogues (see Fig. 1c) are indicated with asterisks.

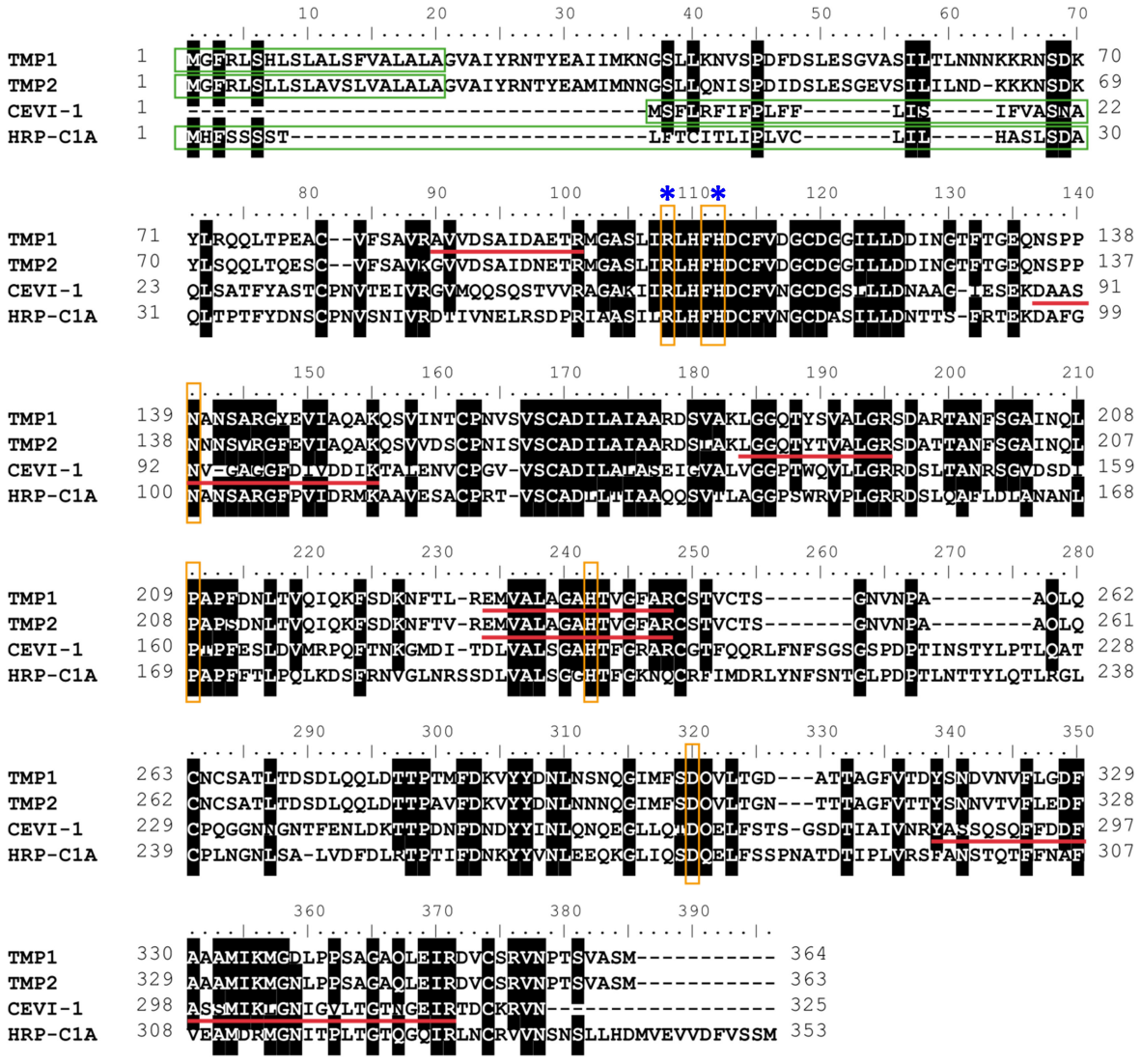


Extended Data Figure 2 | Purification of chemoattractant compounds from tomato root exudate reveals secreted peroxidases. a, Chemotropic growth of germ tubes towards a gradient of tomato root exudate (RE) either untreated (no treat); treated with 1 mg ml^{-1} proteinase K for 30 min at 37 °C (PK); extracted to obtain an ethyl acetate fraction (EAF) and a water fraction (WF); or the WF subjected to centrifugal ultrafiltration with membranes of 10, 30 or 50 kDa molecular weight cut-off to obtain fractions <10, 10–30, 30–50 and >50, respectively (*P = 0.006; **P < 0.0001, versus untreated). **b**, Anion exchange chromatography profile of fraction 30–50 from **a**. Obtained fractions F1–F5 are indicated. **c**, Directed growth of *F. oxysporum* germ tubes towards fractions F1–F5 from **b** (*P < 0.0001, versus H_2O). **d**, SDS–PAGE of biologically active fraction F1 and inactive fraction F5, followed by staining with Coomassie blue. Protein bands present in the active and absent from the inactive fraction (named B1–B5) are indicated by arrowheads. Relative positions of molecular weight markers are indicated on the right. **e**, Directed growth of germ tubes towards the proteins eluted from bands B1–B5. **f**, Peroxidase activity of root exudates obtained from 18 individual tomato plants, indicated as units ml^{-1} per mg fresh root weight. Data are presented as the mean of three technical replicates. Error bars show s.d. **g**, Relationship between peroxidase enzymatic activity of root exudates and elicited chemotropic response. Each empty circle represents a root exudate sample from an individual tomato plant (n = 18). Linear regression (solid line) and 95% mean prediction interval (dashed lines) indicate linear correlation of the two variables (P < 0.001). **h**, Specific inhibitors and oxygen radical scavengers abolish peroxidase enzymatic activity. Activity of 2.5 nM commercial HRP or 100 µl root exudate was measured in the absence (C) or presence of 75 µM of the specific inhibitors thiourea (TU) or SHAM, or 250 µM of the scavenger (+)-sodium L-ascorbate (Asc), and indicated as units ml^{-1}. Data are presented as the mean of three experiments, each with two technical replicates (*P < 0.0002, versus C). **i**, Peroxidase inhibitors and scavengers do not affect chemotropism towards glucose and α-pheromone. Chemotropic growth of germ tubes towards a gradient of glucose or α-pheromone, in the absence (C) or presence of 60 mM SHAM or 160 mM Asc. No significant differences were observed between treated and untreated samples. **a**, **c**, **e**, **g**, **i**, Data represent the mean from two experiments (**a**, **c**, **g**, **i**) or one representative experiment performed twice (**e**). n = 500 germ tubes. Error bars show s.d.

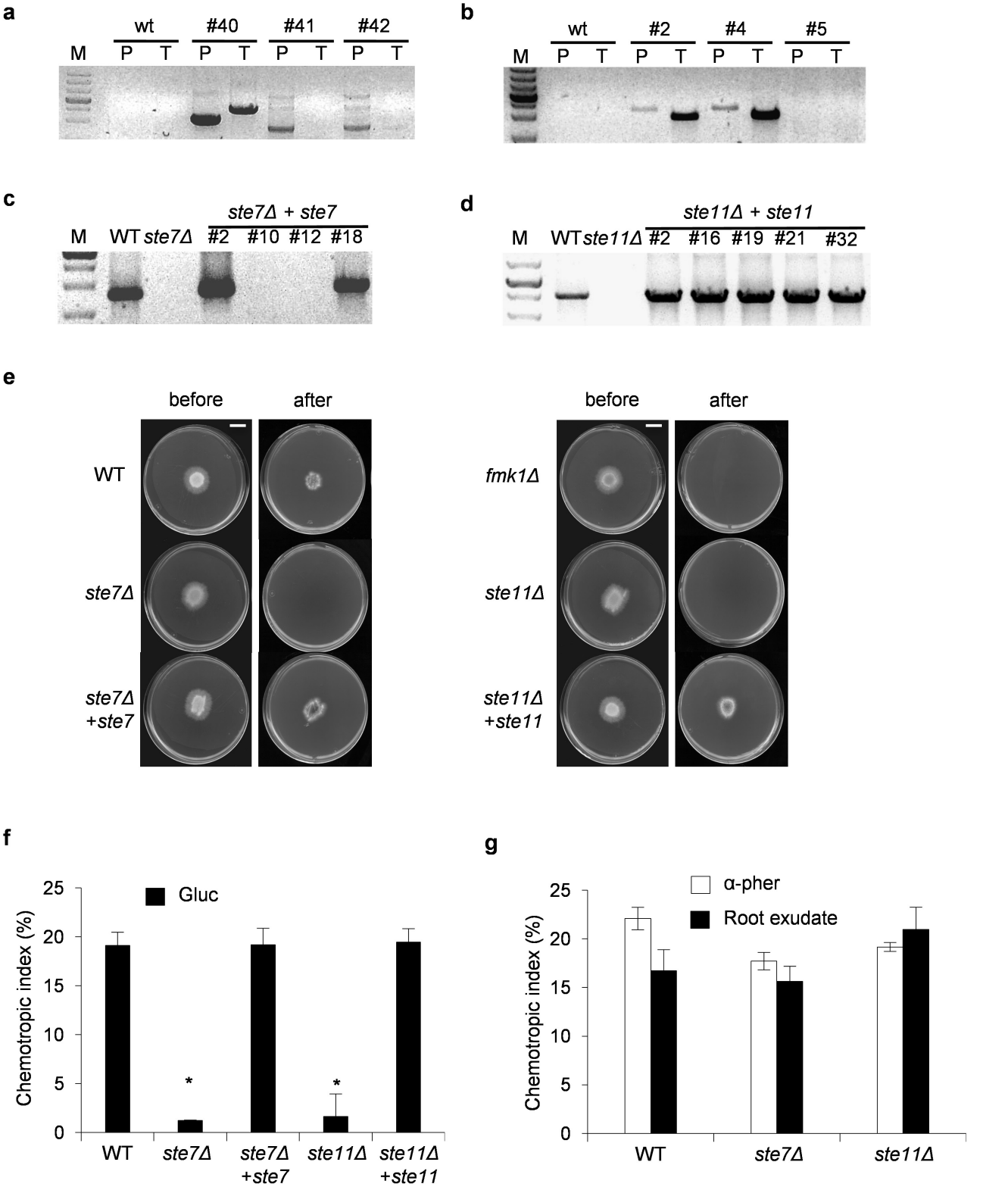
a

Protein name	Protein band	Peptide sequence	Experimental mass (Da)[1]	Theoretical mass (Da)
TMP1	2/3	EMVALAGAHTVGFAR	1545,63/1529,55	1528,78
		AVVDSAIDAETR	1246,51/1246,43	1245,62
CEVI-1	2/3	YASSQSQFFDDFASSMIK	2074,83/2058,67	2057,90
		LGNIGVLTGTNGEIR	1513,79	1512,83
		DAASNVGAGGFDIVDDIK	1763,79	1762,84
TMP2	2	EMVALAGAHTVGFAR	1545,60	1528,78
		LGGQTYTVALGR	1235,40	1234,67

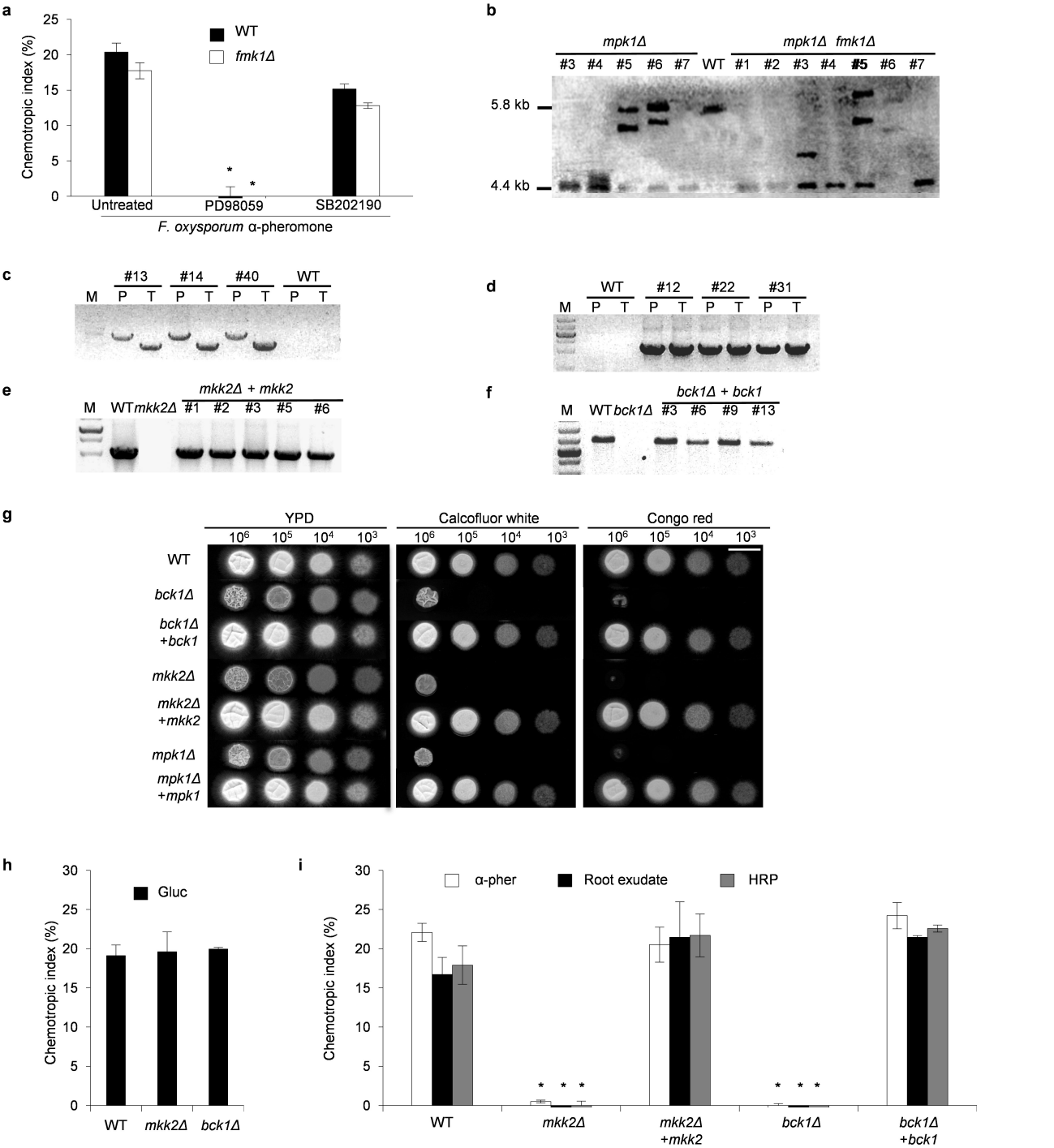
b



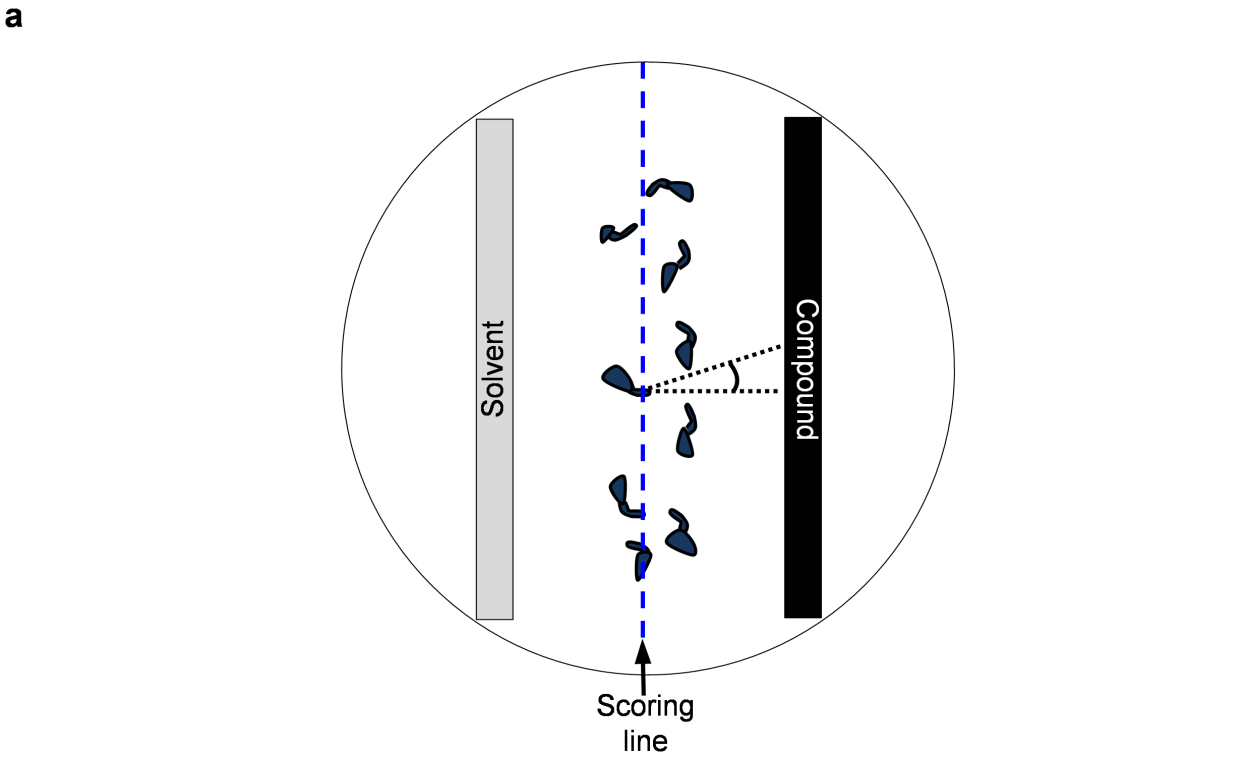
Extended Data Figure 3 | Identification of chemoattractant proteins from tomato root exudate. a, Peptide sequences obtained from protein bands B2 and B3 after in-gel tryptic digestion followed by LC-ESI-MS/MS. Masses were calculated by using monoisotopic masses of the occurring amino acid residues and giving peptide masses as [MH] +. **b**, Amino acid sequence alignment of class III tomato peroxidases TMP1 (P15003), TMP2 (P15004) and CEVI-1 (Q9LWA2), and HRP isoenzyme C (HRP_C1A) (K7ZWW6). Peptides identified in the chemotropically active fraction of tomato root exudate by LC-ESI-MS/MS are underlined in red. Predicted signal peptides are indicated by green boxes. Residues conserved in at least three of the four proteins are shaded in black. Conserved catalytic residues are indicated by orange boxes. Residues Arg 38 and His 42, which were replaced by Ser and Glu, respectively, in the catalytically inactive recombinant TMP2(R38S,H42E) protein (see Fig. 2g, h) are marked with blue asterisks.



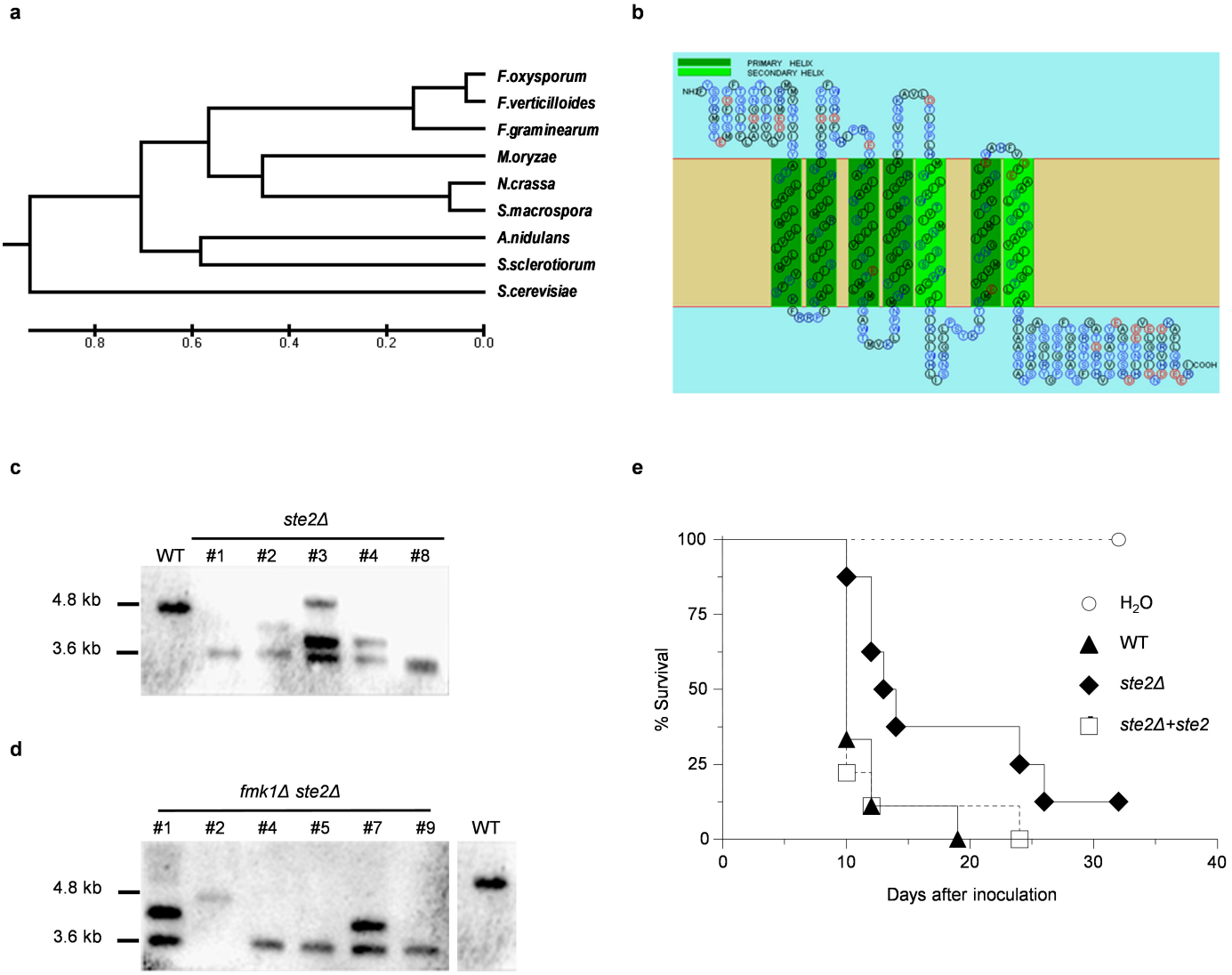
Extended Data Figure 4 | Conserved elements of the invasive growth MAPK cascade are required for chemotropism towards glucose. **a**, **b**, Identification of *ste7*Δ (**a**) and *ste11*Δ (**b**) deletion mutants. Genomic DNA of the wild-type strain (WT) and several independent transformants was used as a template for polymerase chain reaction (PCR) with the primer pairs ste7PF + Hyg-G (P) and ste7TR + Hyg-Y (T) and ste11PF + Hyg-G (P) and ste11TR + Hyg-Y (T), respectively. Presence of an amplification product is consistent with homologous replacement of the target gene. **c**, **d**, Identification of complemented strains obtained from *ste7*Δ and *ste11*Δ mutants. Genomic DNA of independent transformants obtained upon transformation of the indicated mutants with the wild-type *ste7* (**c**) or *ste11* gene (**d**) was used as a template for PCR with primer pairs ste7PFN + ste7GR and ste11PFN + ste11GR, respectively. Presence of an amplification product is consistent with integration of an intact gene copy. **e**, Elements of the Fmk1 MAPK pathway are required for invasive growth through cellophane membranes. Colonies were grown on PDA plates covered with a cellophane membrane for 2 days at 28 °C (before). The cellophane with the fungal colony was removed and plates were incubated for an additional day (after). The experiment was performed twice, each with three plates. Results shown are from one representative experiment. Scale bar, 2 cm. **f**, **g**, Directed growth of germ tubes of the indicated *F. oxysporum* strains towards a gradient of glucose (Gluc) (**f**), α-pheromone or tomato root exudate (**g**) (versus wild type for a given compound, $*P < 0.0001$). **f**, **g**, Data are presented as the mean from two experiments. $n = 500$ germ tubes. Error bars show s.d.



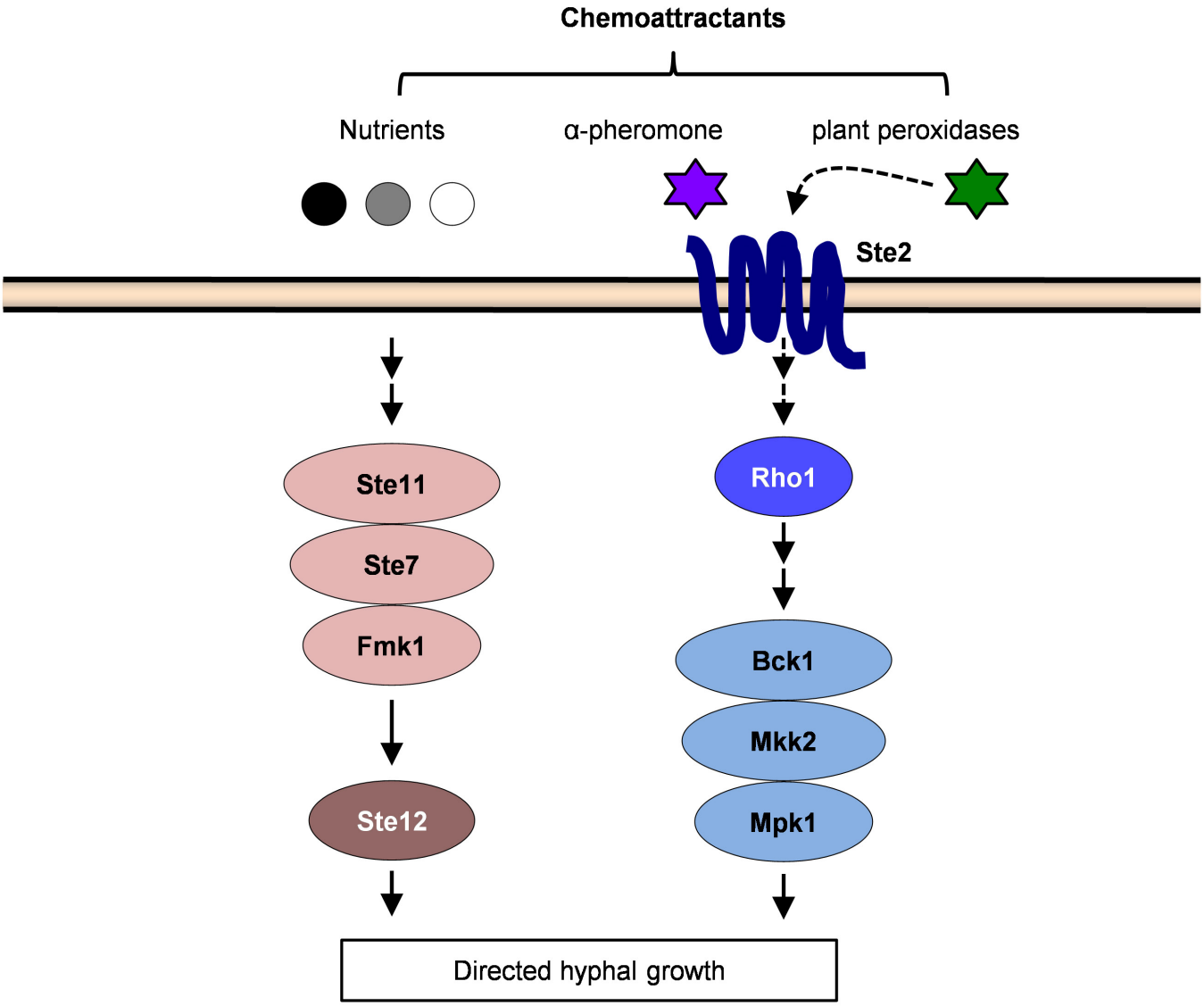
Extended Data Figure 5 | Conserved elements of the CWI MAPK cascade are required for chemotropism towards α-pheromone, root exudate and peroxidase. a, Directed growth of germ tubes of the wild type or the *fmk1*Δ mutant towards a gradient of α-pheromone, in the absence or presence of PD98059 (selective p42/44 (ERK-type) MAPK inhibitor) or SB202190 (selective p38/Hog1 MAPK inhibitor) (versus wild type, $*P < 0.0001$). **b**, Identification of *mpk1*Δ and *fmk1*Δ *mpk1*Δ deletion mutants by Southern blot analysis. Genomic DNA of the wild-type and 11 independent transformants was treated with EcoRI, separated on a 0.7% agarose gel, transferred to a nylon membrane and hybridized with a DNA probe corresponding to the 3′ flanking region of the *mpk1* gene. Transformants #1, #4, #7 (wild-type background) and #1, #2, #4, #7 (*fmk1*Δ background) show a banding pattern consistent with targeted deletion of the *mpk1* gene. **c**, **d**, Identification of *mkk2*Δ (**c**) and *bck1*Δ (**d**) deletion mutants. Genomic DNA of independent transformants was used as template for PCR with the primer pairs mkk2PF + Hyg-G (P) and mkk2TR + Hyg-Y (T), or bck1PF + Hyg-G (P) and bck1TR + Hyg-Y (T), respectively. Presence of an amplification product is consistent with homologous replacement of the target gene. **e**, **f**, Identification of complemented strains obtained from *mkk2*Δ and *bck1*Δ mutants. Genomic DNA of independent transformants obtained after transformation of the indicated mutants with the wild-type *mkk2* (**c**) or *bck1* allele (**d**) was used as a template for PCR with the primer pairs mkk2PFN + mkk2GR, or bck1PFN + bck1GR, respectively. Presence of an amplification product is consistent with integration of an intact gene copy. **g**, Elements of the Mpk1 MAPK pathway are required for the cell wall stress response. Colony phenotypes of the indicated strains grown on yeast peptone dextrose medium (YPD) in the absence or presence of the cell-wall-perturbing compounds Calcofluor white (20 μg ml^{-1}) or Congo red (100 μg ml^{-1}). Plates were spot-inoculated with the indicated amount of microconidia, incubated for 4 days at 28 °C and scanned. The experiment was performed twice, each with three plates. Results shown are from one representative experiment. **h**, **i**, Directed growth of germ tubes of the indicated *F. oxysporum* strains towards a gradient of glucose (Gluc) (**h**), α-pheromone, tomato root exudate or HRP (**i**) (versus wild type for a given compound, $*P < 0.0001$). **a**, **h**, **i**, Data are presented as the mean from two experiments. $n = 500$ germ tubes. Error bars show s.d.



Extended Data Figure 6 | Hyphal tip projection angle assay reveals differential roles of Fmk1 and Mpk1 MAPKs in chemotropism towards glutamate and α-pheromone. a, Schematic representation of the chemotropism plate assay based on measurement of hyphal tip projection angles. **b**, Average cosine of hyphal tip projection angles of the *F. oxysporum* wild-type, *fmk1*Δ or *mpk1*Δ strains towards a gradient of Glu, α-pheromone or the water control. Data are presented as the mean from three experiments. $n = 100$ germ tubes. Bars indicate upper and lower 95% significance limits for cosine means according to a *t*-test. A cosine of 1 means perfect orientation while 0 means random orientation. Chemotropism was considered significant when the lower confidence limit was >0.



Extended Data Figure 7 | Loss of Ste2 negatively affects virulence of *F. oxysporum* on tomato plants. a, Phylogram of Ste2 orthologues from ascomycete fungi. The analysis was conducted using the MEGA5 program. Distances were inferred using the unweighted pair group method with arithmetic mean (UPGMA). **b**. Two-dimensional model of the transmembrane topology of *F. oxysporum* Ste2. The model was generated using the SOSUI software[23]. Amino acid residues in the primary and secondary transmembrane helix are indicated in dark and light green, respectively. Hydrophobic, positively, and negatively charged residues are marked in black, blue, and red, respectively. **c**, **d**, Southern blot analysis to identify *ste2*Δ (**c**) and *fmk1*Δ *ste2*Δ (**d**) deletion mutants. Genomic DNA of wild type and the indicated transformants was treated with EcoRI, separated on a 0.7% agarose gel, transferred to a nylon membrane and hybridized with a DNA probe corresponding to the 5′ flanking sequence of the *ste2* gene. Transformants #1 and #8 in **c** and #4, #5 and #9 in **d** show a banding pattern consistent with targeted deletion of the *ste2* gene by homologous integration of a single construct. **e**, Loss of Ste2 negatively affects virulence of *F. oxysporum* on tomato seedlings. Surface-sterilized tomato seeds (cultivar Monika) were germinated in glass tubes with 4 ml 0.5% water agar containing 2.5×10^6 ml^{-1} microconidia of the indicated *F. oxysporum* strains and incubated at 28 °C under a daily cycle of 15 h light and 9 h dark. Plant survival was recorded for 32 days. Plants inoculated with the *ste2*Δ mutant showed significantly lower mortality than those inoculated with the wild-type and the complemented strain ($P = 0.02$, log-rank test). $n = 40$ plants. Results shown are from one representative experiment. Experiments were performed twice with similar results.



Extended Data Figure 8 | Mechanism of chemotropic signalling in *Fusarium oxysporum*. Chemotropic sensing of nutrients such as glucose or Glu is mediated by the Fmk1 MAPK pathway. Chemotropic sensing of α-pheromone and plant peroxidase-derived signals requires the 7TM-domain receptor Ste2 and the Mpk1 MAPK pathway. Dotted arrows between components denote unknown mechanistic links.

Extended Data Table 1 | *Fusarium oxysporum* strains used in this study

Strain	Genotype	Gene function	Reference
FGSC 4287	wild type		(ref. 16)
4287 *GFP*	*PgpdA-GFP*; *HYG*	Green Fluorescent Protein	(ref. 16)
fmk1Δ	*fmk1::PHLEO*	MAPK	(ref. 16)
fmk1Δ + fmk1	*fmk1::PHLEO; fmk1::HYG*	MAPK	(ref. 16)
ste12Δ	*ste12::HYG*	Homeodomain transcription factor	(ref. 29)
ste12Δ + ste12	*ste12::HYG; ste12::PHLEO*	Homeodomain transcription factor	(ref. 29)
rho1Δ	*rho1::HYG*	Rho-type GTPase	(ref. 19)
rho1Δ + rho1	*rho1::HYG; rho1::PHLEO*	Rho-type GTPase	(ref. 19)
ste2Δ	*ste2::HYG*	GPCR	This study
ste2Δ + ste2	*ste2::HYG; ste2::PHLEO*	GPCR	This study
fmk1Δ ste2Δ	*fmk1::PHLEO; ste2::HYG*	MAPK/GPCR	This study
mpk1Δ	*mpk1::HYG*	MAPK	This study
mpk1Δ + mpk1	*mpk1::HYG; mpk1::PHLEO*	MAPK	This study
fmk1Δ mpk1Δ	*fmk1::PHLEO; mpk1::HYG*	MAPK / MAPK	This study
bck1Δ	*bck1::HYG*	MAPKKK	This study
bck1Δ + bck1	*bck1::HYG; bck1::PHLEO*	MAPKKK	This study
mkk2Δ	*mkk2::HYG*	MAPKK	This study
mkk2Δ + mkk2	*mkk2::HYG;mkk2::PHLEO*	MAPKK	This study
ste11Δ	*ste11::HYG*	MAPKKK	This study
ste11Δ + ste11	*ste11::HYG; ste11::PHLEO*	MAPKKK	This study
ste7Δ	*ste7::HYG*	MAPKK	This study
ste7 + ste7Δ	*ste7::HYG; ste7::PHLEO*	MAPKK	This study

Ref. 29 is cited in this Table.

Extended Data Table 2 | Oligonucleotides used in this study

Primer	Sequence	Use
gpdA15B	CGAGACCTAATACAGCCCCT	*hygB* cassette
trpter8B	GGATCCAAACAAGTGTACCTGTGCATTC	
Hyg-G	CGTTGCAAGACCTGCCTGAA	
Hyg-Y	GGATGCCTCCGCTCGAAGTA	
Ste2PFO	GCAGGCACAAAGAACAGCAAT	*ste2* knockout/ complementation
Ste2PFN	GTGGCAGAGGAGAGAGCTATAG	
Ste2PFN2	ATTACACCAGCAGTGTTTGCC	
Ste2PR	TAAAGATTGGAAGTGAAAGGGG	
Ste2PRGPDA15B	*TGGTCGTTGTAGGGGCTGTATTAGGTCTCG*(A)TAAAGATTGGAAGTGAAAGGGG*	
Ste2TR	TCAACATCAACAAGCGAAAGAG	
Ste2TRN	AACTTAGGGGCTCTGAGGATG	
Ste2TFTrpter8B	*TTTACCCAGAATGCACAGGTACACTTGTTT*(A)GACCAAAACAAAACTTCTAGCG*	
Ste2PFO2	ACCTGGATACACGAACGATAC	
Mpk1TF1	*TGGTCGTTGTAGGGGCTGTATTAGGTCTCG*(A)CAGTATTTCCCTTCAGCCAAC*	*mpk1* knockout/ complementation
Mpk1TR1	TCAAGACCAATGTACCTACGG	
Mpk1TR2	CTTTTGGACGAACTGTGAACC	
Mpk1PF1	TGGAGAAGAGTAAATGGACGG	
Mpk1PF2	AGGGAAACGAGGTAGGTTACA	
Mpk1PR1	*TTTACCCAGAATGCACAGGTACACTTGTTT*(A)CTGGTGATGTGGCTGATTTGT*	
Mpk1PF0	TCCACAGACTACAGAAGAACG	
Mpk1-R	TCTCCTAGAGGCATCCAGTCC	
STE11 PF	TAGGTGATTAGACGTGGGAAG	*ste11* knockout/ complementation
STE11 PFN	GGTCTAGGCTCACTTTGTTTC	
STE11 PR	*TTTACCCAGAATGCACAGGTACACTTGTTT*(A)CATTGTGGGCTGAGAAGGAAC*	
STE11 TF	*TGGTCGTTGTAGGGGCTGTATTAGGTCTCG*(A)TTGACCACAACCTACGACCTA*	
STE11 TRN	CTTCTGATATGCCGATGGAAC	
STE11 TR	CATCAGTCCTTCTCAATCCAG	
STE11 GR	ATGTTCAGGGATTGTAGGAGC	
STE7 PF	CCTGAGCCGAGTATGGAATTG	*ste7* knockout/ complementation
STE7 PFN	GTCCCCTTATGGCGAATGAAT	
STE7 PR	*TTTACCCAGAATGCACAGGTACACTTGTTT*(A)GATAGAATTGACAAGCTCGCC*	
STE7 TF	*TGGTCGTTGTAGGGGCTGTATTAGGTCTCG*(A)TACCGTCTTCAAATCCCAAGG*	
STE7 TRN	CAGTGGCTTCGTTAATCAGTC	
STE7 TR	AATGGAAGAGAGTGGAAGAGG	
STE7 GR	TAAATGATCTTCAGGGTTAGGC	
BCK1 PF	GAGACTTTTGGAATGGAGAGG	*bck1* knockout/ complementation
BCK1 PFN	GGTAGATTGAGTTACGTCTGG	
BCK1 PR	*TTTACCCAGAATGCACAGGTACACTTGTTT*(A)TCTTGAGGCTGAGATTGAGAC*	
BCK1 TF	*TGGTCGTTGTAGGGGCTGTATTAGGTCTCG*(A)GTCTGGGTTGTGTAGTCCTG*	
BCK1 TRN	TGGTGATGTTCGTCAAGAGATA	
BCK1 TR	GTTTCCTTGTTGCCTCGATCT	
BCK1 GR	ATCCAGAATACCGAACCTTGC	
MKK2 PF	TAGCTTTGGATTGCGGTTGGA	*mkk2* knockout/ complementation
MKK2 PFN	ACGAGAATGACGATGTGTGTG	
MKK2 PR	*TTTACCCAGAATGCACAGGTACACTTGTTT*(A)GGGTAGTGGAGTTGAATCAGA*	
MKK2 TF	*TGGTCGTTGTAGGGGCTGTATTAGGTCTCG*(A)TTTCTTTTGTCTGGGGTTGGG*	
MKK2 TRN	CCGAGAATAGCATCTTCAGAC	
MKK2 TR	CAGATTGCTCGTTTCCTCAAG	
MKK2 GR	GCTCGTCTGTTGGGTTGTTTT	
Tap2_for1	CTCTGTTTCTTCCAAATAGAC	*tmp2* cloning
Tap1/2_rev	CTCGAGTCACATAGAAGCCACAGAAG	
Tmp2FPCR_for	ATTAGTCTACATTTCGAGGACTGC	*tmp2* mutagenesis
Tmp2FPCR_rev	GTCCTCGAAATGTAGACTAATGAG	
Cevi1_for1	CATATGCAATTAAGTGCAACATTTTACG	*cevi-1* cloning
Cevi1_rev	CTCGAGCTAATCAACTAATTAACCCTCT	

*The sequence shown in italics corresponds to the complementary region of the gpdA15B (Ste2PRGPDA15B, Mpk1TF1, STE11TF, STE7TF, BCK1TF and MKK2TF) or trpter8B (Ste2TFTrpter8B, Mpk1PR1, STE11PR, STE7PR, BCK1PR and MKK2PR) primers.

LETTER

doi:10.1038/nature16064

Epithelial–to–mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer

Xiaofeng Zheng[1]*, Julienne L. Carstens[1]*, Jiha Kim[1], Matthew Scheible[1], Judith Kaye[1], Hikaru Sugimoto[1], Chia–Chin Wu[2], Valerie S. LeBleu[1] & Raghu Kalluri[1,3,4]

Diagnosis of pancreatic ductal adenocarcinoma (PDAC) is associated with a dismal prognosis despite current best therapies; therefore new treatment strategies are urgently required. Numerous studies have suggested that epithelial-to-mesenchymal transition (EMT) contributes to early-stage dissemination of cancer cells and is pivotal for invasion and metastasis of PDAC[1–4]. EMT is associated with phenotypic conversion of epithelial cells into mesenchymal-like cells in cell culture conditions, although such defined mesenchymal conversion (with spindle-shaped morphology) of epithelial cells *in vivo* is rare, with quasi-mesenchymal phenotypes occasionally observed in the tumour (partial EMT)[5,6]. Most studies exploring the functional role of EMT in tumours have depended on cell-culture-induced loss-of-function and gain-of-function experiments involving EMT-inducing transcription factors such as Twist, Snail and Zeb1 (refs 2,3,7–10). Therefore, the functional contribution of EMT to invasion and metastasis remains unclear[4,6], and genetically engineered mouse models to address a causal connection are lacking. Here we functionally probe the role of EMT in PDAC by generating mouse models of PDAC with deletion of Snail or Twist, two key transcription factors responsible for EMT. EMT suppression in the primary tumour does not alter the emergence of invasive PDAC, systemic dissemination or metastasis. Suppression of EMT leads to an increase in cancer cell proliferation with enhanced expression of nucleoside transporters in tumours, contributing to enhanced sensitivity to gemcitabine treatment and increased overall survival of mice. Collectively, our study suggests that Snail- or Twist-induced EMT is not rate-limiting for invasion and metastasis, but highlights the importance of combining EMT inhibition with chemotherapy for the treatment of pancreatic cancer.

We crossed *Twist1*$^{loxP/loxP}$ (*Twist1*$^{L/L}$) or *Snai1*$^{loxP/loxP}$ (*Snai1*$^{L/L}$) mice with *Pdx1-cre;LSL-Kras*^{G12D}*;P53*$^{R172H/+}$ (KPC) to generate the *Pdx1-cre;LSL-Kras*^{G12D}*;P53*$^{R172H/+}$*;Twist1*$^{L/L}$ (KPC;Twist^{cKO}) and the *Pdx1-cre;LSL-Kras*^{G12D}*;P53*$^{R172H/+}$*;Snai1*$^{L/L}$ (KPC;Snail^{cKO}) mice, respectively. The resultant progeny were born in an expected Mendelian ratio, without overt phenotypic findings other than the anticipated emergence of spontaneous pancreatic cancer (Extended Data Fig. 1a). Genetic deletion of *Snai1* or *Twist1* did not significantly delay pancreatic tumorigenesis, alter tumour histopathology features or local invasion (Fig. 1a–c and Extended Data Table 1). KPC;Twist^{cKO} and KPC;Snail^{cKO} mice displayed similar tumour burden compared to KPC control mice (Extended Data Fig. 1b) and insignificant differences in overall survival (Fig. 1d). Loss of *Twist1* or *Snai1* expression in the pancreas epithelium was confirmed by *in situ* hybridization coupled with CK8 epithelial immunolabelling (Fig. 1e and Extended Data Fig. 1c) as well as immunolabelling for Twist and Snail (Extended Data Fig. 1d). Significant suppression of EMT was noted (Fig. 1f, g and Extended Data Fig. 1e, f). Lineage tracing (Fig. 1f and Extended Data Fig. 1e) and immunolabelling of the primary tumour (Fig. 1g) showed a significant decrease in the frequency of epithelial cells with expression of the mesenchymal marker αSMA (EMT^{+} cells) and a decrease in expression of the EMT-inducing transcription factor Zeb1 (Fig. 1h). Global gene expression profiling of tumours revealed a decrease in expression of EMT-associated genes (including *Snai1* and *Twist1*) in KPC;Snail^{cKO} and KPC;Twist^{cKO} mice compared to KPC control (Extended Data Fig. 1f). Loss of Snail and Twist enhanced E-cadherin expression and suppressed Zeb2 and Sox4 expression in cancer cells (Extended Data Fig. 2a–c). Snai2 (Slug) expression was restricted to early pancreatic intraepithelial neoplasia (PanIN) lesions in all the experimental groups with no observed expression in advanced tumours and was significantly reduced in KPC;Snail^{cKO} and KPC;Twist^{cKO} mice compared to KPC control mice (Extended Data Fig. 2d).

While desmoplasia, including extracellular matrix (ECM) and myofibroblasts content (Fig. 1i and Extended Data Fig. 2e, f), tumour vessel density (Extended Data Fig. 2g), intratumoural hypoxia (Extended Data Fig. 2h), CD3^{+} T-cell infiltration (Extended Data Fig. 2i), and cancer cell apoptosis was unaffected with Twist/Snail deletion in KPC tumours (Fig. 2a), the proliferation of cancer cells in mice with suppressed EMT was significantly increased (Fig. 2b), as shown previously in mouse models of breast cancers[11–13]. Immunostaining experiments further revealed that EMT^{+} cancer cells are largely Ki67^{-} (Extended Data Fig. 3a). Altogether, these data suggest that EMT driven by Twist/Snail transcription factors is dispensable for initiation and progression of primary pancreatic cancer.

Next, we investigated whether suppression of EMT impacts invasion and metastasis. The number of YFP^{+} circulating tumour cells from lineage-traced KPC and KPC;Twist^{cKO} was found to be unchanged (Fig. 2c and Extended Data Fig. 3b), and expression of cancer-cell-specific Kras^{G12D} mRNA in the blood from KPC, KPC;Twist^{cKO} and KPC;Snail^{cKO} mice was unaffected (Fig. 2d), suggesting that suppression of EMT in pancreatic tumours does not impact the rate of systemic dissemination of cancer cells. Extensive histopathological analyses, coupled with CK19 or YFP immunostaining of distant metastatic target organs, namely the liver, lung and spleen, indicated a similar frequency of metastasis in EMT-suppressed tumours when compared to control tumours (Fig. 2e, Extended Data Fig. 3c and Extended Data Tables 1 and 2). The metastases were negative for Twist, Snail, Zeb1 and αSMA, with the exception of a few KPC metastatic cells that expressed αSMA or Zeb1 (Extended Data Fig. 3d–f),

[1]Department of Cancer Biology, Metastasis Research Center, University of Texas MD Anderson Cancer Center, Houston, Texas 77054, USA. [2]Department of Genomic Medicine, University of Texas MD Anderson Cancer Center, Houston, Texas 77054, USA. [3]Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, Texas 77030, USA. [4]Department of Bioengineering, Rice University, Houston, Texas 77030, USA.
*These authors contributed equally to this work.

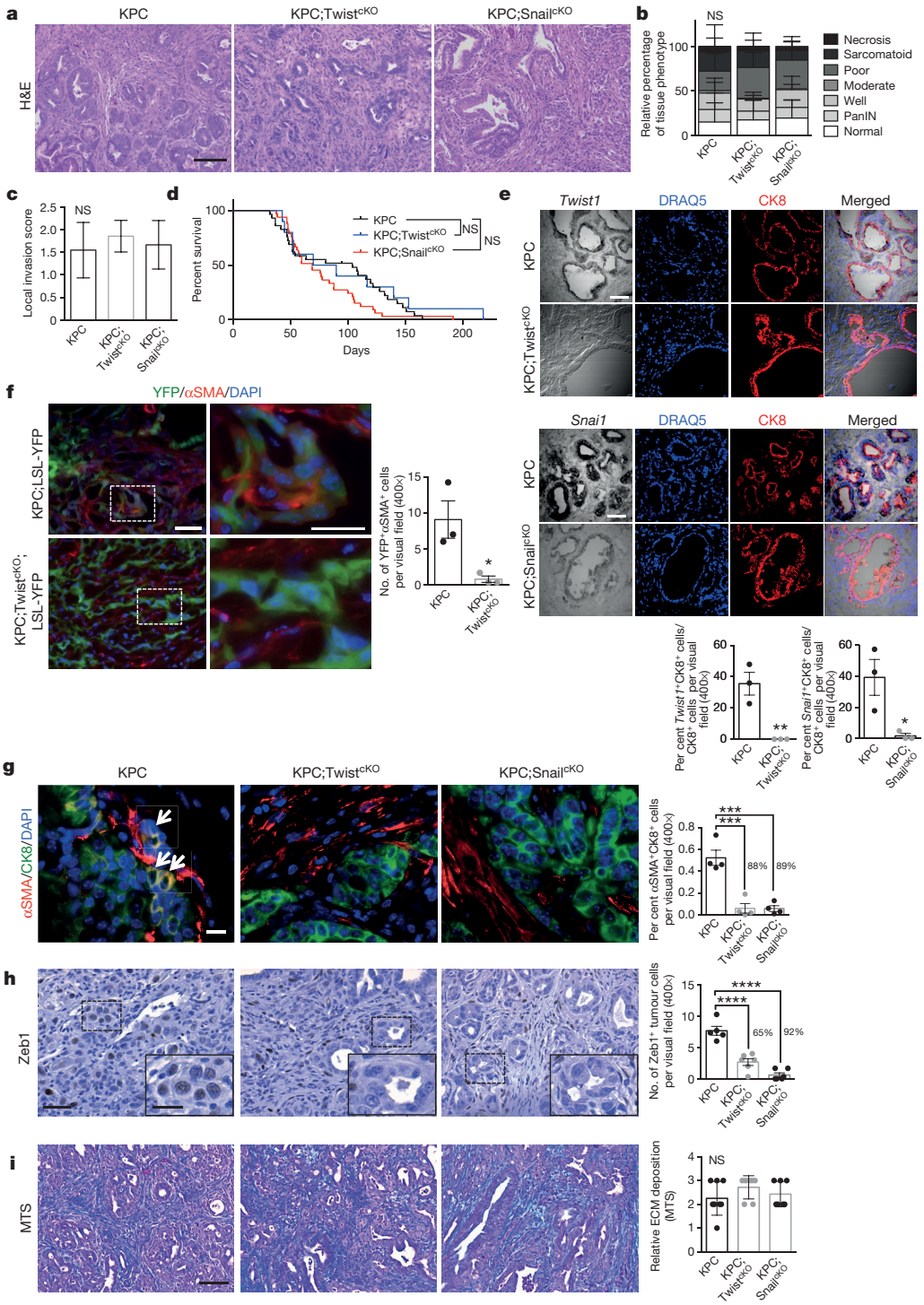
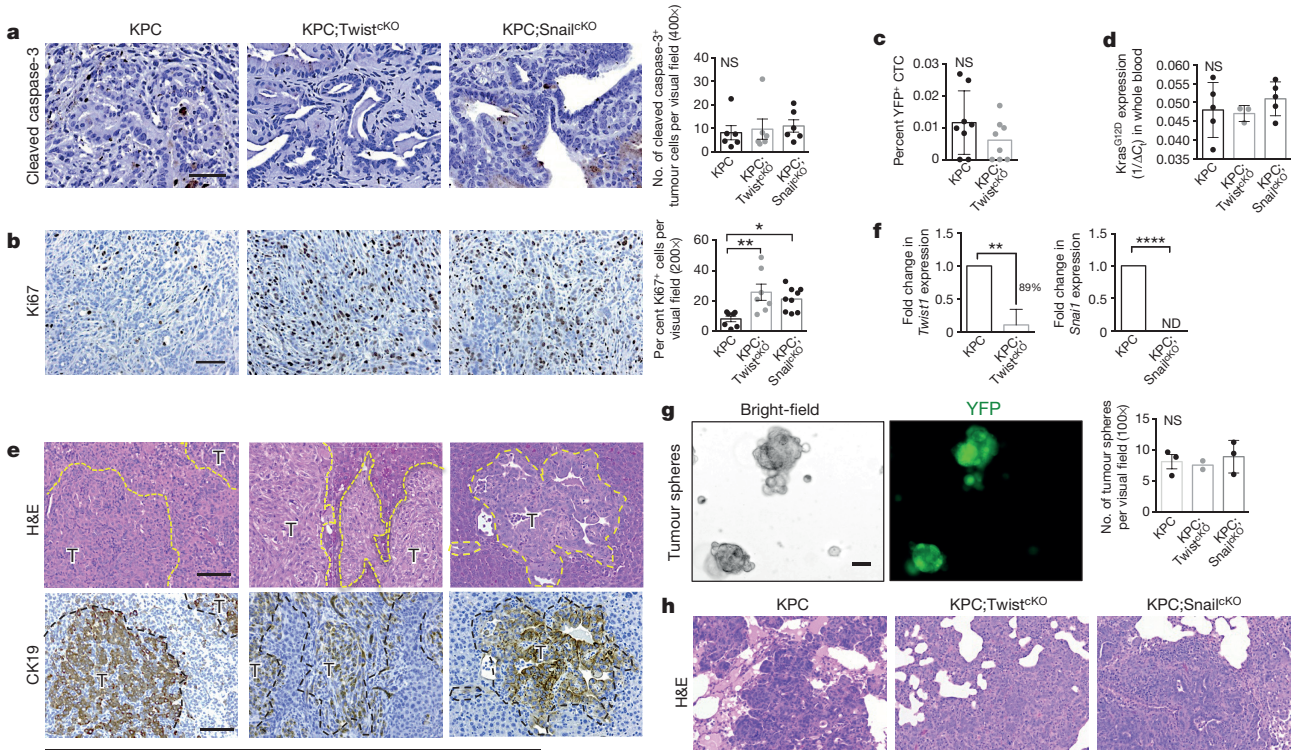


Figure 1 | EMT inhibition does not alter primary tumour progression. **a**, Representative haematoxylin and eosin (H&E)-stained primary tumours (scale bar, 100 μm). **b**, Relative percentages of each primary tumour histological tissue phenotype. $n = 31$ (KPC), 14 (KPC;Twist^{cKO}) and 30 (KPC;Snail^{cKO}) mice; error bars represent s.d. **c**, Local invasiveness $n = 31$ (KPC), 14 (KPC;Twist^{cKO}) and 30 (KPC;Snail^{cKO}) mice; error bars represent s.d. **d**, Overall survival $n = 29$ (KPC), 12 (KPC;Twist^{cKO}) and 33 (KPC;Snail^{cKO}) mice. **e**, *Twist1* or *Snai1 in situ* hybridization (black) with CK8 (red) immunolabelling in primary tumours ($n = 3$ mice for all groups; scale bar, 50 μm). Relative percentages of *Twist1*$^+$CK8$^+$ or *Snai1*$^+$CK8$^+$ double-positive cells are shown below (two-tailed *t*-test). **f**, αSMA immunolabelling in YFP lineage-traced primary tumours ($n = 3$ mice for both groups; scale bar, 50 μm; two-tailed *t*-test). **g**, αSMA (red), CK8 (green) and DAPI (blue) immunolabelling in primary tumours; white arrows indicate double-positive cells ($n = 4$ mice for all groups; scale bar, 20 μm). **h**, Zeb1 immunolabelling ($n = 5$ (KPC), 6 (KPC;Twist^{cKO}) and 6 (KPC;Snail^{cKO}) mice; scale bar, 50 μm; inset scale bar, 20 μm). **i**, Masson's trichrome stain (MTS) ($n = 8$ (KPC), 7 (KPC;Twist^{cKO}) and 7 (KPC;Snail^{cKO}) mice; scale bar, 100 μm; error bars represent s.d.). Unless otherwise indicated error bars represent s.e.m., percentages represent per cent change from control and significance was determined by one-way ANOVA. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$; NS, not significant.

while being positive for E-cadherin and Ki-67 (Extended Data Fig. 3g, h). The proliferation rate of cancer cells in the metastases was similar in KPC, KPC;Snail^{cKO} and KPC;Twist^{cKO} mice (Extended Data Fig. 3h). Collectively, the results indicated that the deletion of *Twist1* or *Snai1* in genetically engineered mouse models of PDAC did not reduce metastatic disease.



	KPC	KPC;Twist^{cKO}	KPC;Snail^{cKO}
Liver metastasis	11/31	6/14	13/30
Lung metastasis	11/31	4/14	9/30
Spleen invasion	2/30	2/14	5/29
Any metastasis	17/31	8/14	18/30

No significant differences

	KPC		KPC;Twist^{cKO}		KPC;Snail^{cKO}		
Lung colonization	3/5	4/5	9/11	4/4	4/4	5/5	5/5
Liver colonization	1/5	0/5	2/11	0/4	0/4	1/5	1/5
Spleen colonization	0/5	0/5	0/11	0/4	0/4	0/5	0/5

No significant differences

Figure 2 | EMT inhibition does not alter invasion and metastasis. **a**, **b**, Primary tumour immunolabelling for cleaved caspase-3 (**a**; $n = 6$ mice for all groups; scale bar, 50 μm) and Ki67 (**b**; $n = 7$ (KPC), 7 (KPC;Twist^{cKO}) and 9 (KPC;Snail^{cKO}) mice; scale bar, 100 μm). **c**, Percentage of YFP$^+$ circulating tumour cells (CTCs) ($n = 8$ mice for both groups; two-tailed *t*-test; error bars represent s.d.). **d**, Kras^{G12D} expression in whole blood cell pellets ($n = 5$ (KPC), 3 (KPC;Twist^{cKO}) and 5 (KPC;Snail^{cKO}) mice; error bars represent s.d.). **e**, Haematoxylin and eosin staining and CK19 immunolabelling of metastatic liver nodules. Metastatic tumour nodules (T) outlined by a dotted line (scale bar, 100 μm). A table presenting the number of positive tissues out of total tissues examined is shown below (χ^2 analysis). **f**, Expression analysis of *Twist1* and *Snai1* in cultured primary tumour cell lines ($n = 4$ (KPC) and 5 (KPC;Twist^{cKO}) individual cell lines (*Twist1*) or 4 (KPC) and 6 (KPC;Snail^{cKO}) individual cell lines (*Snai1*); one-tailed *t*-test of ΔC_t; error bars represent s.d.). **g**, Bright-field or YFP images and quantification of sphere number in cultured tumour cell lines ($n = 3$ (KPC), 2 (KPC;Twist^{cKO}) and 3 (KPC;Snail^{cKO}) individual cell lines; scale bar, 50 μm). **h**, Haematoxylin and eosin images (scale bar, 100 μm) of colonized lungs from intravenously injected cultured primary tumour cell lines KPC ($n = 5$ (cell line 1) and 5 (cell line 2) mice injected) and KPC;Twist^{cKO} ($n = 11$ (cell line 1) and 4 (cell line 2) mice injected) and KPC;Snail^{cKO} ($n = 4$ (cell line 1), 5 (cell line 2) and 5 (cell line 3) mice injected). A table presenting the number of colonized tissues out of total tissues examined is shown below (χ^2 analysis). Unless otherwise indicated error bars represent s.e.m and significance was determined by one-way ANOVA. $^*P < 0.05$, $^{**}P < 0.01$, $^{****}P < 0.0001$; NS, not significant; ND, not detected.

To evaluate whether cancer cells from the pancreas with and without EMT program differentially benefited from impaired proliferation to form secondary tumours, we isolated cancer cells from KPC, KPC;Twist^{cKO} and KPC;Snail^{cKO} mice to assay their organ colonization potential. *Twist1* was significantly reduced and *Snai1* expression was undetectable in cancer cells isolated from Twist- and Snail-deleted tumours, respectively (Fig. 2f). Short-term potential to form tumour spheres (associated with putative cancer stem phenotype) appeared similar in Twist^{cKO} and Snail^{cKO} KPC cells when compared to control KPC cells (Fig. 2g)[3,8,14–16]. Lung colonization frequencies following i.v. injection of KPC cancer cells (Twist- or Snail-deleted) were similar to the control KPC cancer cells (Fig. 2h). These results suggest that a favoured epithelial phenotype of cancer cells (via suppression of EMT) did not impact the capacity to form tumour spheres or their ability for organ colonization[17].

Cancer cell EMT is associated with gemcitabine drug resistance in PDAC patients and in the orthotopic mouse models of PDAC[1,2,8,9,18–23]. Moreover, enhanced frequency of EMT$^+$ cancer cells in pancreatic tumours is associated with poor survival[24,25]. To determine whether EMT suppression enhances PDAC sensitivity to gemcitabine chemotherapy, we tested the gemcitabine sensitivity of cancer cells with suppressed EMT in KPC mice. Equilibrative nucleoside transporter (ENT1) and concentrating nucleoside transporter (Cnt3) were significantly upregulated in cancer cells lacking Snail and Twist, while ENT2 expression was unchanged (Fig. 3a–c). KPC, KPC;Snail^{cKO} and KPC;Twist^{cKO} mice were treated with gemcitabine and tumour burden was monitored by MRI (Extended Data Table 3). Tumour progression was suppressed in KPC;Snail^{cKO} and KPC;Twist^{cKO} mice when compared to treated KPC control mice (Fig. 3d). KPC;Snail^{cKO} and KPC;Twist^{cKO} mice treated with gemcitabine showed improved histopathology and increased survival (Fig. 3e–g).

Cancer cells isolated from the tumours of KPC;Snail^{cKO} and KPC;Twist^{cKO} mice showed epithelial morphology (Extended Data Fig. 4a) and reduced expression of mesenchymal genes compared to KPC cancer cell lines (Extended Data Fig. 4b). However, in tissue culture conditions (2D culture on plastic), equilibrative nucleoside transporters (ENT1/ENT2/ENT3) showed similar expression patterns (Extended Data Fig. 4b) and expression of concentrating nucleoside transporters (Cnt1/Cnt3) was not detected (data not shown). Increased proliferation of KPC;Snail^{cKO} and KPC;Twist^{cKO} cancer cells compared to KPC control cells (Extended Data Fig. 4c) probably accounted for

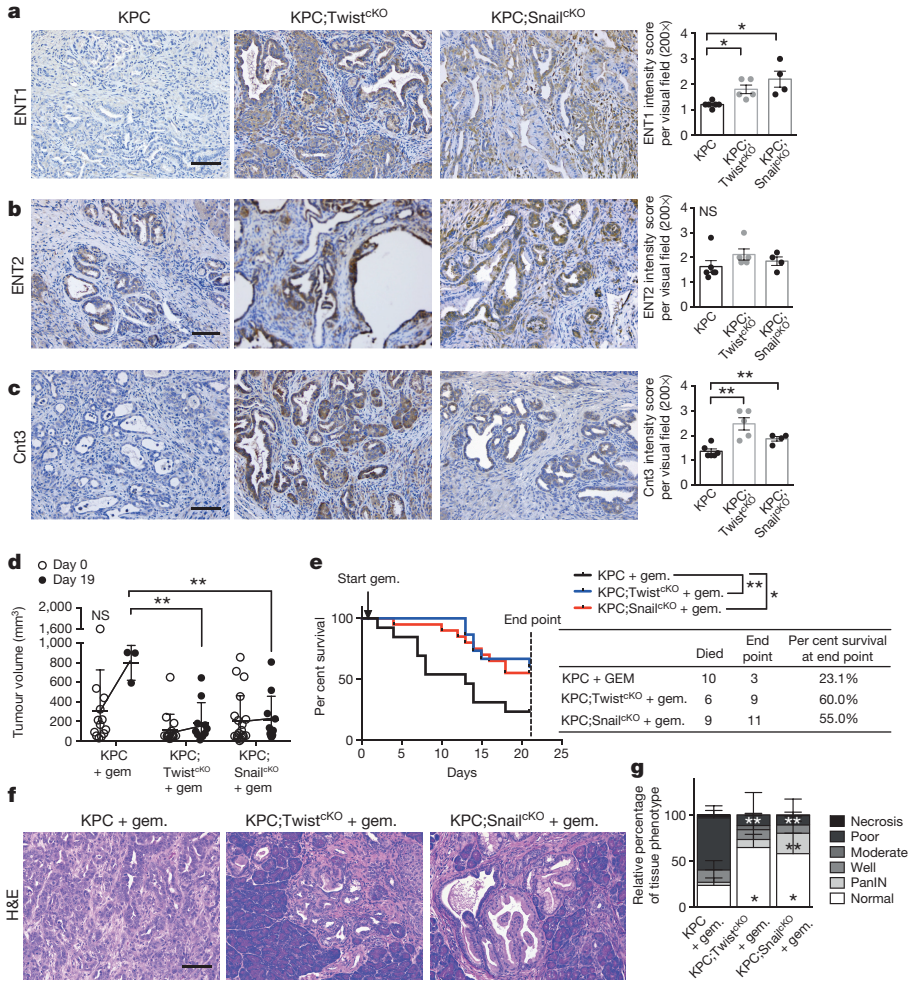


Figure 3 | EMT inhibition sensitizes tumours to gemcitabine in KPC mice. a–c, Primary tumour immunolabelling for (ENT1 (**a**), ENT2 (**b**) and Cnt3 (**c**) ($n = 6$ (KPC), 5 (KPC;Twist^{cKO}) and 4 (KPC;Snail^{cKO}) mice; scale bar, 100 μm; error bars represent s.e.m., two-tailed t-test). **d**, MRI tumour volumes of KPC plus gemcitabine (+ gem.) ($n = 13$ mice, 10 died before day 19), KPC; Twist^{cKO} + gem. ($n = 15$ mice, 5 died before day 19) and KPC; Snail^{cKO} + gem. ($n = 20$ mice, 9 died before day 19). One-way ANOVA comparing mean tumour volumes on day 0 and day 19, error bars represent s.d. **e**, Survival on gemcitabine treatment to end point (day 21). **f**, Haematoxylin and eosin-stained primary tumours (scale bar, 100 μm). **g**, Relative percentages of each histological tissue phenotype of end-point mice ($n = 3$ (KPC + gem.), 9 (KPC; Twist^{cKO} + gem.) and 11 (KPC; Snail^{cKO} + gem.) mice; error bars represent s.d.; two-tailed t-test). $^*P < 0.05$, $^{**}P < 0.01$; NS, not significant.

the increased sensitivity to gemcitabine and erlotinib in this setting (Extended Data Fig. 4d).

Next, we crossed the *Snai1*$^{L/L}$ to the PDAC mouse model, *Ptf1a (P48)-cre;LSL-Kras*^{G12D}*;Tgfbr2*$^{L/L}$ (KTC) to generate *Ptf1a (P48)-cre;LSL-Kras*^{G12D}*;Tgfbr2*$^{L/L}$*;Snai1*$^{L/L}$ (KTC; Snail^{cKO}). The KTC model offers a reliable and penetrant disease progression rate with a consistent timeline of death due to PDAC. Similar to the KPC;Snail^{cKO} mice, KTC;Snail^{cKO} deletion exhibited suppression of EMT but did not affect primary tumour histopathology, lifespan, local invasion, desmoplasia or frequency of apoptosis (Fig. 4f and Extended Data Figs 5a–e and 6a). KTC;Snail^{cKO} mice presented with significantly reduced Zeb1 expression in cancer cells but enhanced expression of Cnt3, ENT2 and proliferation (Extended Data Fig. 5e). ENT1 expression was unchanged in KTC;Snail^{cKO} mice compared to KTC mice (Extended Data Fig. 6a). KTC;Snail^{cKO} mice demonstrated enhanced response to gemcitabine therapy, with significant normal parenchymal area and reduced tumour tissue (Fig. 4a–c). Gemcitabine therapy in KTC;Snail^{cKO} mice reduced tumour burden (Fig. 4d) and significantly improved overall survival (Fig. 4e) when compared to gemcitabine-treated control KTC mice. Gemcitabine therapy specifically increased cancer cell apoptosis and removed enhanced proliferation observed in EMT-suppressed tumours (Fig. 4g and Extended Data Fig. 5e), without impacting the desmoplastic reaction (Extended Data Fig. 6b). Overall, these results suggest an enhanced sensitivity of EMT-suppressed cancer cells to gemcitabine. Both ENT2 and Cnt3 were upregulated in EMT-suppressed tumours (Fig. 4g). These data support a possible mechanistic connection between EMT and resistance to chemotherapy in PDAC.

Collectively, our studies provide a comprehensive functional analysis of EMT in PDAC progression and metastasis. Absence of either *Twist1* or *Snai1* did not alter cancer progression or the capacity for local invasion or metastasis to lung and liver in genetically engineered mouse models of PDAC. Metastasis occurs despite a significant loss of EMT with either the deletion of Snail or Twist, and in both settings, Zeb1, Sox4, Slug and Zeb2 are also significantly suppressed. Nevertheless, it is possible that other EMT-inducing factors may compensate for the loss of Snail or Twist to induce invasion and metastasis. While Pdx1 is expressed during the development of the pancreas (in early pancreatic buds and all three major lineages of the pancreas: ductal, acinar and β-islets), its expression is largely repressed in the adult exocrine pancreas[26,27]. Therefore, deletion of Snail or Twist occurs at the embryonic stage and mice are born normal and exhibit normal pancreas histology before the onset of cancer. The mice with Snail or Twist deletion develop PanIN lesions at the same frequency as the control mice. One could argue that suppression of EMT starting from the inception of cancer could have launched compensatory mechanisms to overcome EMT-dependent invasion and metastasis.

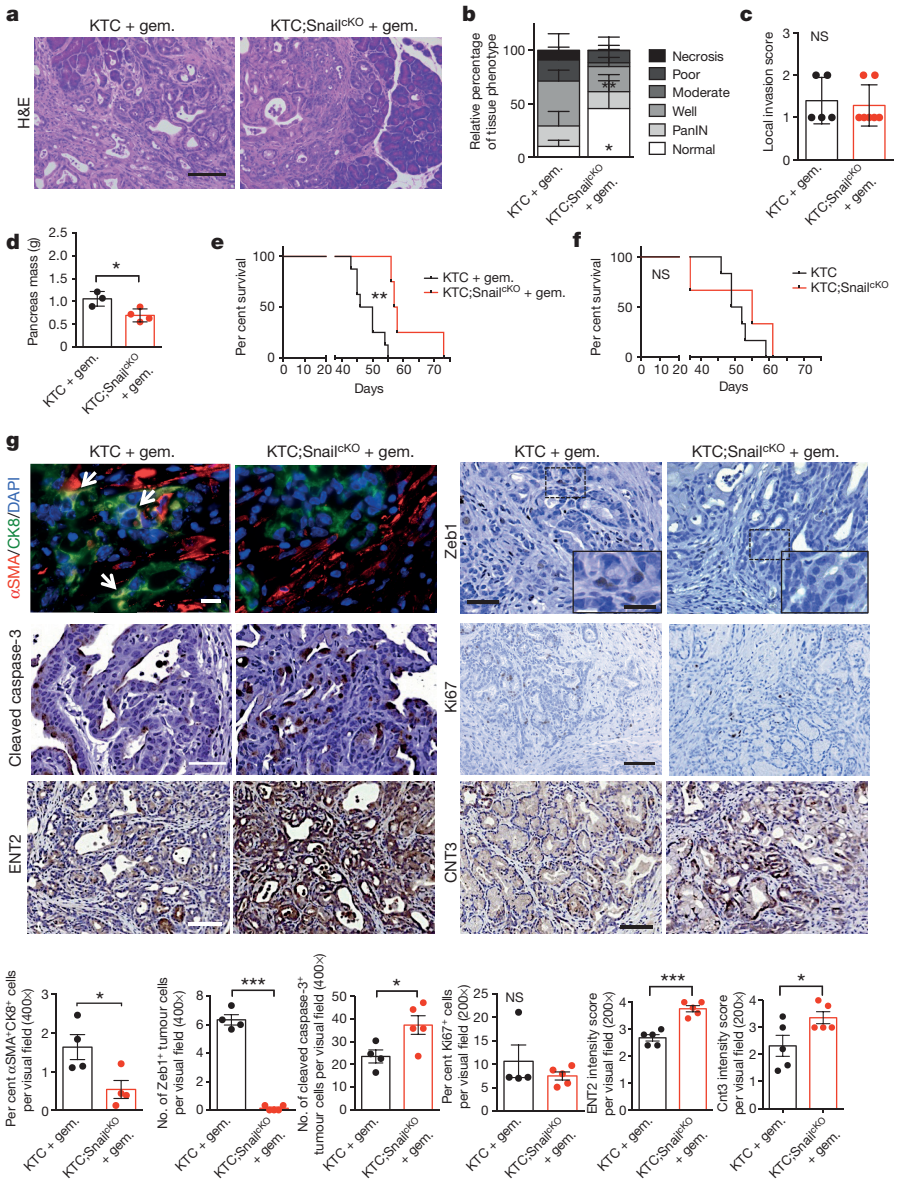


Figure 4 | EMT inhibition sensitizes tumours to gemcitabine in KTC mice. a, **b**, Haematoxylin and eosin-stained primary tumour (scale bar, 100 μm) and relative percentage of each histological tissue phenotype in KTC + gem. ($n = 5$) and KTC;Snail^{cKO} + gem. ($n = 7$) mice (error bars represent s.d.). **c**, Local invasiveness ($n = 5$ (KTC + gem.) and 7 (KTC;Snail^{cKO} + gem.) mice; error bars represent s.d.). **d**, Pancreatic mass ($n = 3$ (KTC + gem.) and 4 (KTC;Snail^{cKO} + gem.) mice; error bars represent s.d.). **e**, Overall survival of KTC + gem. ($n = 8$) and KTC;Snail^{cKO} + gem. ($n = 4$) mice. **f**, Overall survival of KTC ($n = 6$) and KTC;Snail^{cKO} ($n = 3$) mice. **g**, αSMA (red), CK8 (green) and DAPI (blue) staining of primary tumours; white arrows indicate double-positive cells ($n = 4$ mice for both groups; scale bar, 20 μm), and immunolabelling for Zeb1 ($n = 4$ (KTC + gem.) and 5 (KTC;Snail^{cKO} + gem.) mice; scale bar, 50 μm; inset scale bar, 20 μm), cleaved caspase-3 ($n = 4$ (KTC + gem.) and 5 (KTC;Snail^{cKO} + gem.) mice; scale bar, 50 μm), Ki67 ($n = 4$ (KTC + gem.) and 5 (KTC;Snail^{cKO} + gem.) mice; scale bar, 100 μm), ENT2 ($n = 5$ mice for both groups; scale bar, 100 μm), and Cnt3 ($n = 5$ mice for both groups; scale bar, 100 μm). Unless otherwise indicated error bars represent s.e.m. and significance was determined by two-tailed *t*-tests. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$; NS, not significant.

However, such compensation is not observed with respect to chemoresistance, and previous studies have demonstrated that EMT and cancer cell dissemination are observed even before PDAC lesions are detected in KPC mice[4].

Our study demonstrates that EMT results in suppression of cancer cell proliferation and suppression of drug transporter and concentrating proteins, therefore inadvertently protecting EMT$^+$ cells from anti-proliferative drugs such as gemcitabine. The correlation of decreased survival of pancreatic cancer patients with increased EMT is probably due to their impaired capacity to respond to gemcitabine and chemotherapeutics, which is a standard of care for most patients[28,29]. A compromised response to chemotherapy probably also explains higher metastatic disease in association with decreased survival of patients with enhanced EMT signatures. Collectively, our study offers the opportunity to evaluate the potential of targeting EMT to enhance efficacy of chemotherapy and targeted therapies[30].

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 10 June; accepted 8 October 2015.
Published online 11 November 2015.

1. Hotz, B. *et al.* Epithelial to mesenchymal transition: expression of the regulators snail, slug, and twist in pancreatic cancer. *Clin. Cancer Res.* **13,** 4769–4776 (2007).
2. Arumugam, T. *et al.* Epithelial to mesenchymal transition contributes to drug resistance in pancreatic cancer. *Cancer Res.* **69,** 5820–5828 (2009).

3. Taube, J. H. *et al.* Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. *Proc. Natl Acad. Sci. USA* **107,** 15449–15454 (2010).
4. Rhim, A. D. *et al.* EMT and dissemination precede pancreatic tumor formation. *Cell* **148,** 349–361 (2012).
5. Kalluri, R. & Weinberg, R. A. The basics of epithelial–mesenchymal transition. *J. Clin. Invest.* **119,** 1420–1428 (2009).
6. McDonald, O. G., Maitra, A. & Hruban, R. H. Human correlates of provocative questions in pancreatic pathology. *Adv. Anat. Pathol.* **19,** 351–362 (2012).
7. Guaita, S. *et al.* Snail induction of epithelial to mesenchymal transition in tumor cells is accompanied by MUC1 repression and ZEB1 expression. *J. Biol. Chem.* **277,** 39209–39216 (2002).
8. Wellner, U. *et al.* The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs. *Nature Cell Biol.* **11,** 1487–1495 (2009).
9. Zhang, K. *et al.* Knockdown of snail sensitizes pancreatic cancer cells to chemotherapeutic agents and irradiation. *Int. J. Mol. Sci.* **11,** 4891–4904 (2010).
10. Tsai, J. H., Donaher, J. L., Murphy, D. A., Chau, S. & Yang, J. Spatiotemporal regulation of epithelial–mesenchymal transition is essential for squamous cell carcinoma metastasis. *Cancer Cell* **22,** 725–736 (2012).
11. Stockinger, A., Eger, A., Wolf, J., Beug, H. & Foisner, R. E-cadherin regulates cell growth by modulating proliferation-dependent β-catenin transcriptional activity. *J. Cell Biol.* **154,** 1185–1196 (2001).
12. Muraoka-Cook, R. S., Dumont, N. & Arteaga, C. L. Dual role of transforming growth factor beta in mammary tumorigenesis and metastatic progression. *Clin. Cancer Res.* **11,** 937s–943s (2005).
13. Hugo, H. J. *et al.* Direct repression of MYB by ZEB1 suppresses proliferation and epithelial gene expression during epithelial-to-mesenchymal transition of breast cancer cells. *Breast Cancer Res.* **15,** R113 (2013).
14. Mani, S. A. *et al.* The epithelial–mesenchymal transition generates cells with properties of stem cells. *Cell* **133,** 704–715 (2008).
15. Liu, H. *et al.* Cancer stem cells from human breast tumors are involved in spontaneous metastases in orthotopic mouse models. *Proc. Natl Acad. Sci. USA* **107,** 18115–18120 (2010).
16. Wang, Z. *et al.* Activated K-Ras and INK4a/Arf deficiency promote aggressiveness of pancreatic cancer by induction of EMT consistent with cancer stem cell phenotype. *J. Cell. Physiol.* **228,** 556–562 (2013).
17. Yang, J. *et al.* Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell* **117,** 927–939 (2004).
18. Vega, S. *et al.* Snail blocks the cell cycle and confers resistance to cell death. *Genes Dev.* **18,** 1131–1143 (2004).
19. Shah, A. N. *et al.* Development and characterization of gemcitabine-resistant pancreatic tumor cells. *Ann. Surg. Oncol.* **14,** 3629–3637 (2007).
20. Yin, T. *et al.* Expression of snail in pancreatic cancer promotes metastasis and chemoresistance. *J. Surg. Res.* **141,** 196–203 (2007).
21. Wang, Z. *et al.* Acquisition of epithelial–mesenchymal transition phenotype of gemcitabine-resistant pancreatic cancer cells is linked with activation of the notch signaling pathway. *Cancer Res.* **69,** 2400–2407 (2009).
22. Alagesan, B. *et al.* Combined MEK and PI3K inhibition in a mouse model of pancreatic cancer. *Clin. Cancer Res.* **21,** 396–404 (2015).
23. Cursons, J. *et al.* Stimulus-dependent differences in signalling regulate epithelial-mesenchymal plasticity and change the effects of drugs in breast cancer cell lines. *Cell Commun. Signal.* **13,** 26 (2015).
24. Javle, M. M. *et al.* Epithelial–mesenchymal transition (EMT) and activated extracellular signal-regulated kinase (p-Erk) in surgically resected pancreatic cancer. *Ann. Surg. Oncol.* **14,** 3527–3533 (2007).
25. Masugi, Y. *et al.* Solitary cell infiltration is a novel indicator of poor prognosis and epithelial-mesenchymal transition in pancreatic cancer. *Hum. Pathol.* **41,** 1061–1068 (2010).
26. Park, J. Y. *et al.* Pdx1 expression in pancreatic precursor lesions and neoplasms. *Appl. Immunohistochem. Mol. Morphol.* **19,** 444–449 (2011).
27. Offield, M. F. *et al.* PDX-1 is required for pancreatic outgrowth and differentiation of the rostral duodenum. *Development* **122,** 983–995 (1996).
28. Hidalgo, M. Pancreatic cancer. *N. Engl. J. Med.* **362,** 1605–1617 (2010).
29. Kleger, A., Perkhofer, L. & Seufferlein, T. Smarter drugs emerging in pancreatic cancer therapy. *Ann. Oncol.* **25,** 1260–1270 (2014).
30. Gore, A. J., Deitz, S. L., Palam, L. R., Craven, K. E. & Korc, M. Pancreatic cancer-associated retinoblastoma 1 dysfunction enables TGF-β to promote proliferation. *J. Clin. Invest.* **124,** 338–352 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements We wish to thank D. Lundy, S. Yang, Z. Xiao, R. Deliz-Aguirre, T. Miyake and S. Lovisa for technical support and K. M. Ramirez and R. Jewell in the South Campus Flow Cytometry Core Laboratory of MD Anderson Cancer Center for flow cytometry cell sorting and analyses (partly supported by NCI grant no. P30CA16672). We also wish to thank E. Chang for scanning slides of histopathological specimens. This study was primarily supported by the Cancer Prevention and Research Institute of Texas. The research in the LeBleu laboratory is supported by UT MDACC Khalifa Bin Zayed Al Nahya Foundation.

Author Contributions R.K. conceptually designed the strategy for this study and provided intellectual input. V.S.L. helped design experimental strategy, provided intellectual input, supervised the studies, performed immunohistochemistry and culture experiments, generated the figures and wrote the manuscript. X.Z. performed experiments to generate the genetically engineered mouse models and helped characterize the mouse phenotype, performed culture experiments, collected the tissue for analysis and contributed to the manuscript writing. J.L.C. characterized the mouse phenotype, analysed the data related to the genetically engineered mouse models, collected data, generated the figures and helped with manuscript writing and editing. H.S. performed experiments with mice and injected cancer cells and helped collect tissue, J.Ki., M.S., J.Ka., and C.-C.W. performed experiments and collected data. The data was analysed by J.L.C, V.S.L., X.Z., J.Ki, and C.-C.W.

Author Information Gene expression microarray data have been deposited in the Gene Expression Omnibus under accession number GSE66981. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper.
Correspondence and requests for materials should be addressed to R.K. (rkalluri@mdanderson.org).

METHODS

Mice. Characterization of disease progression and genotyping for the *Pdx1-cre;LSL-Kras^{G12D};P53$^{R172H/+}$* (herein referred to as KPC) and *Ptf1a (P48)-cre;LSL-Kras^{G12D};Tgfbr2$^{L/L}$* (herein referred to as KTC) mice were previously described[31–33]. These mice were bred to *Snai1$^{L/L}$* (herein referred to as Snail^{cKO}), *Twist1$^{L/L}$* (herein referred to as Twist^{cKO}), and *R26-LSL-EYFP*[33]. Snail^{cKO} mice were kindly provided by S. J. Weiss. Twist^{cKO} mice were kindly provided by R. R. Behringer via the Mutant Mouse Regional Resource Center (MMRRC) repository. The resulting progeny were referred to as KPC, KPC;Snail^{cKO}, KPC;Twist^{cKO}, KTC and KTC;Snail^{cKO} mice and were maintained on a mixed genetic background. Both males and females were used indiscriminately. Mice were given gemcitabine (G-4177, LC Laboratories) via intraperitoneal injection (i.p.) every other day at 50 mg kg^{-1} of body weight. Hypoxyprobe was injected in a subset of mice i.p. at 60 mg kg^{-1} of body weight 30 min before euthanasia. For *in vivo* colonization assays, one million KPC, KPC;Twist^{cKO} and KPC;Snail^{cKO} tumour cells in 100 μl of PBS were injected intravenously via the retro-orbital venous sinus. Four to eleven mice were injected per cell line. All mice were euthanized at 15 days post injection. All mice were housed under standard housing conditions at MD Anderson Cancer Center (MDACC) animal facilities, and all animal procedures were reviewed and approved by the MDACC Institutional Animal Care and Use Committee. Tumour growth met the standard of a diameter less than or equal to 1.5 cm. Investigators were not blinded to group allocation but were blinded for the assessment of the phenotypic outcome by histological analyses. No statistical methods were used to predetermine sample size and the experiments were not randomized.

Histology and histopathology. Histology, histopathological scoring, Masson's trichrome staining (MTS), and Picrosirius Red have been previously described[19,33]. Formalin-fixed tissues were embedded in paraffin and sectioned at 5 μm thickness. MTS was performed using Gomori's Trichome Stain Kit (38016SS2, Leica Biosystems). Picrosirius red staining for collagen was performed using 0.1% picrosirius red (Direct Red80; Sigma) and counterstained with Weigert's haematoxylin. Sections were also stained with haematoxylin and eosin (H&E). Histopathological measurements were assessed by scoring H&E-stained tumours for relative percentages of each histopathological phenotype: normal (non-neoplastic), PanIN, well-differentiated PDAC, moderately-differentiated PDAC, poorly-differentiated PDAC, sarcomatoid carcinoma, or necrosis. When tumour histology was missing or of poor quality, the mice were excluded from primary tumour histological analysis and this was determined blinded from genotype information. A histological invasion score of the tumour cells into the surrounding stroma was scored on a scale of 0 to 2, with 0 indicating no invasion and 2 indicating high invasion, where invasion is defined as tumour cell dissemination throughout the stroma away from clearly defined epithelial 'nests'. Microscopic metastases were observed in H&E-stained tissue sections of the liver, lung and spleen. Positivity (one or more lesions in a tissue) was confirmed using CK19 and YFP immunohistochemistry. This data has been presented as a contingency table (Fig. 2e) and represented as the number of positive tissues out of the number of tissues scored. The 'Any' metastasis score is the number of mice positive for a secondary lesion found anywhere throughout the body out of the total number of mice scored.

Immunohistochemistry and Immunofluorescence. Tissues were fixed in 10% formalin overnight, dehydrated, and embedded in paraffin and 5-μm-thick sections were then processed for analyses. Immunohistochemical analysis was performed as described[33]. Heat-mediated antigen retrieval in 1 mM EDTA + 0.05% Tween20 (pH 8.0) for one hour (pressure cooker) was performed for Snail and Twist, 10 mM citrate buffer, pH 6.0, was used for one hour (microwave) for Ki67 or 10 min for all other antibodies. Primary antibodies are as follows: αSMA (M0851, DAKO, 1:400 or ab5694, Abcam, 1:400), cleaved caspase-3 (9661, Cell Signaling, 1:200), CD3 (A0452, DAKO, 1:200), CD31 (Dia310M, DiaNova, 1:10), CK8 (TROMA-1, Developmental Studies Hybridoma Bank, 1:50), CK19 (ab52625, Abcam, 1:100), Cnt3 (HPA023311, Sigma-Aldrich, 1:400), ENT1 (LS-B3385, LifeSpan Bio., 1:100), E-cadherin (3195S, Cell Signaling, 1:400), ENT2 (ab48595, Abcam, 1:200), Ki67 (RM-9106, Thermo Scientific, 1:400), Slug (9585, Cell Signaling, 1:200), Snail (ab180714, Abcam, 1:100), Sox4 (ab86809, Abcam, 1:200), Twist (ab50581, Abcam, 1:100), YFP (ab13970, Abcam, 1:1000), Zeb1 (NBP1-05987, Novus, 1:500), and Zeb2 (NBP1-82991, Novus, 1:100). Sections for pimonidazole adduct (HPI Inc., 1:50) or αSMA immunohistochemistry staining were blocked with M.O.M. kit (Vector Laboratories, West Grove, PA) and developed by DAB according to the manufacturer's recommendations. Alternatively, for immunofluorescence, sections were dual-labelled using secondary antibodies conjugated to Alexa Fluor 488 or 594 or tyramide signal amplification (TSA, PerkinElmer) conjugated to FITC. Lineage-traced (YFP-positive) EMT analysis was performed on 8-μm-thick O.C.T. medium (TissueTek)-embedded frozen sections. Sections were stained for αSMA (ab5694, Abcam, 1:400) followed by Alexa Fluor 680 conjugated secondary antibody. Bright-field imagery was obtained on a Leica DM1000 light microscope or the Perkin Elmer 3DHistotech Slide Scanner. Fluorescence imagery was obtained on a Zeiss Axio Imager.M2 or the Perkin Elmer Vectra Multispectral imaging platform. The images were quantified for per cent positive area using NIH ImageJ analysis software (αSMA, Pimonidazole, Slug, and CD31), per cent positive cells using InForm analysis software (Ki67 and CD3), or scored for intensity either positive or negative (αSMA/CK8 dual staining, αSMA, CK19, YFP, Zeb1, Zeb2, Sox4, E-cadherin and cleaved caspase-3) or on a scale of 1–3 (E-cadherin) or 1–4 (ENT1, ENT2 and Cnt3).

***In situ* hybridization.** *In situ* hybridization (ISH) was performed on frozen tumour sections as previously described[34]. In brief, 10-μm-thick sections were hybridized with antisense probes to *Twist1* and *Snai1* overnight at 65 °C. After hybridization, sections were washed and incubated with AP-conjugated sheep anti-DIG antibody (1:2,000; Roche) for 90 min at room temperature. After three washes, sections were incubated in BM Purple (Roche) until positive staining was seen. Digoxigenin-labelled *in situ* riboprobes were generated with an *in vitro* transcription method (Promega and Roche) using a PCR template. The following primers were used to generate the template PCR product. *Twist1*, forward, 5′-CGGCCAGGTACATCGACTTC-3′; reverse, 5′-TAATACGACTCACTATAGGGAGATTTAAAAGTGTGCCCCACGC-3′; *Snai1*, forward, 5′-CAACCGTGCTTTTGCTGAC-3′; reverse, 5′-TAATACGACTCACTATAGGGAGACCTTTAAAATGTAAACATCTTTCTCC-3′.

Gene expression profiling. Total RNA was isolated from tumours of KPC control, KPC;Twist^{cKO} and KPC;Snail^{cKO} mice ($n = 3$ in each group) by TRIzol (15596026, Life Technologies) and submitted to the Microarray Core Facility at MD Anderson Cancer Center. Gene expression analysis was performed using MouseWG-6 v2.0 Gene Expression BeadChip (Illumina). The Limma package from R Bioconductor[35] was used for quantile normalization of expression arrays and to analyse differentially expressed genes between cKO and control sample groups. Gene expression microarray data have been deposited in GEO (Accession number GSE66981). Genes upregulated in cells acquiring an EMT program were expected to be downregulated in the Twist^{cKO} and Snail^{cKO} tumours compared to control tumours.

CTC assays. Blood (200 μl) was collected from KPC;LSL-YFP and KPC;Twist^{cKO};LSL-YFP (ROSA-LSL-YFP lineage tracing of cancer cells) mice and incubated with 10 ml of ACK lysis buffer (A1049201, Gibco) at room temperature to lyse red blood cells. Cell pellets were resuspended in 2% FBS containing PBS and analysed for the number of YFP$^+$ cells by flow cytometry (BD LSRFortessa X-20 Cell Analyzer). The data was expressed as the percentage of YFP$^+$ cells from gated cells, with 100,000 cells analysed at the time of acquisition. Whole blood cell pellets were also assayed for the expression of *Kras^{G12D}* transcripts, using quantitative real-time PCR analyses (described below).

Primary pancreatic adenocarcinoma cell culture and analyses. Derivation of primary PDAC cell lines were performed as previously described[36]. Fresh tumours were minced with sterile razor blades, digested with dispase II (17105041, Gibco, 4 mg ml^{-1})/collagenase IV (17104019, Gibco, 4 mg ml^{-1})/RPMI for 1 h at 37 °C, filtered by a 70 μm cell strainer, resuspended in RPMI/20%FBS and then seeded on collagen I-coated plates (087747, Fisher Scientific). Cells were maintained in RPMI medium with 20% FBS and 1% penicillin, streptomycin and amphotericin B (PSA) antibiotic mixture. Cancer cells were further purified by FACS based on YFP or E-cadherin expression (anti-E-cadherin antibody, 50-3249-82, eBioscience, 1:100). The sorted cells, using BD FACSAria^{TM} II sorter (South Campus Flow Cytometry Core Lab of MD Anderson Cancer Center) were subsequently expanded *in vitro*. All studies were performed on cells cultivated less than 30 passages. As these are primary cell lines, no further authentication methods were applicable and no mycoplasma tests were performed.

MTT and drug sensitivity assays. MTT assay was performed to detect cell proliferation and viability by using Thiazolyl Blue Tetrazolium Bromide (MTT, M2128, Sigma) following the manufacturer's recommendations with an incubation of two hours at 37 °C. For the drug treatment studies, a cell line derived from each of the KPC, KPC;Snail^{cKO} and KPC;Twist^{cKO} mice was treated with 20 μM gemcitabine (G-4177, LC Laboratories) or 100 μM erlotinib (5083S, NEB) for 48 h. The relative cell viability was detected using MTT assay with a cell line derived from each of the KPC, KPC;Snail^{cKO} and KPC;Twist^{cKO} mice. *n* is defined as the number of biological replicates of a single cell line. Control conditions included 1% DMSO vehicle for erlotinib. The relative absorbance was normalized and control (time 0 h or vehicle-treated) arbitrarily set to 1 or 100% for absorbance or drug survival, respectively.

Quantitative real-time PCR analyses (qPCR). RNA was extracted from whole blood cell pellets following ACK lysis using the PicoPure Extraction kit as directed (KIT0214, Arcturus), or from cultured primary pancreatic adenocarcinoma cells using TRIzol (15596026, Life Technologies). cDNA was synthetized using TaqMan Reverse Transcription Reagents (N8080234,

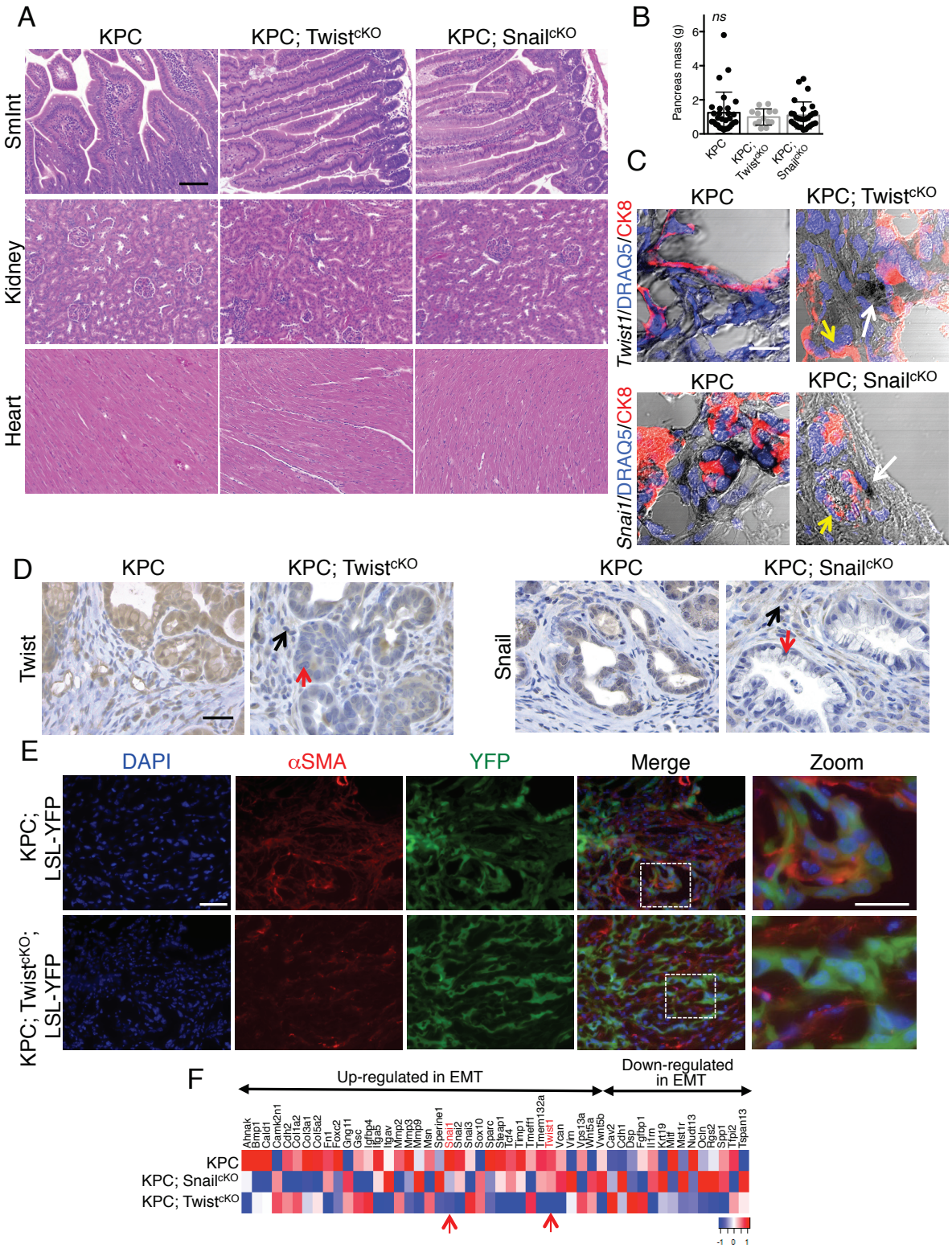
Applied Biosystems) or High Capacity cDNA Reverse Transcription Kit (4368814, Applied Biosystems). Primers for $Kras^{G12D}$ recombination are: $Kras^{G12D}$, forward , 5′-ACTTGTGGTGGTTGGAGCAGC-3′; reverse, 5′-TAGGGTCATACTCATCCACAA-3′. $1/\Delta C_t$ values are presented to show $Kras^{G12D}$ expression in indicated experimental groups, statistical analyses were performed on ΔC_t. Primer sequences for EMT-related genes are listed in Supplementary Table 1, GAPDH was used as an internal control. The data are presented as the relative fold change and statistical analyses were performed on ΔC_t.

Tumour sphere assay. Tumour sphere assays were performed as previously described[33]. Two million cultured primary tumour cells were plated in a low-adherence 100-mm dish (FB0875713, Fisherbrand) with 1% FBS, Dulbecco's modified Eagle's medium, and penicillin/streptomycin/amphotericin. Cells were incubated for 7 days and formed spheres were counted at 100× magnification. Three, two and three cell lines were analysed for KPC control, KPC;$Twist^{cKO}$ and KPC;$Snail^{cKO}$ groups, respectively, five field of views per cell line were quantified.

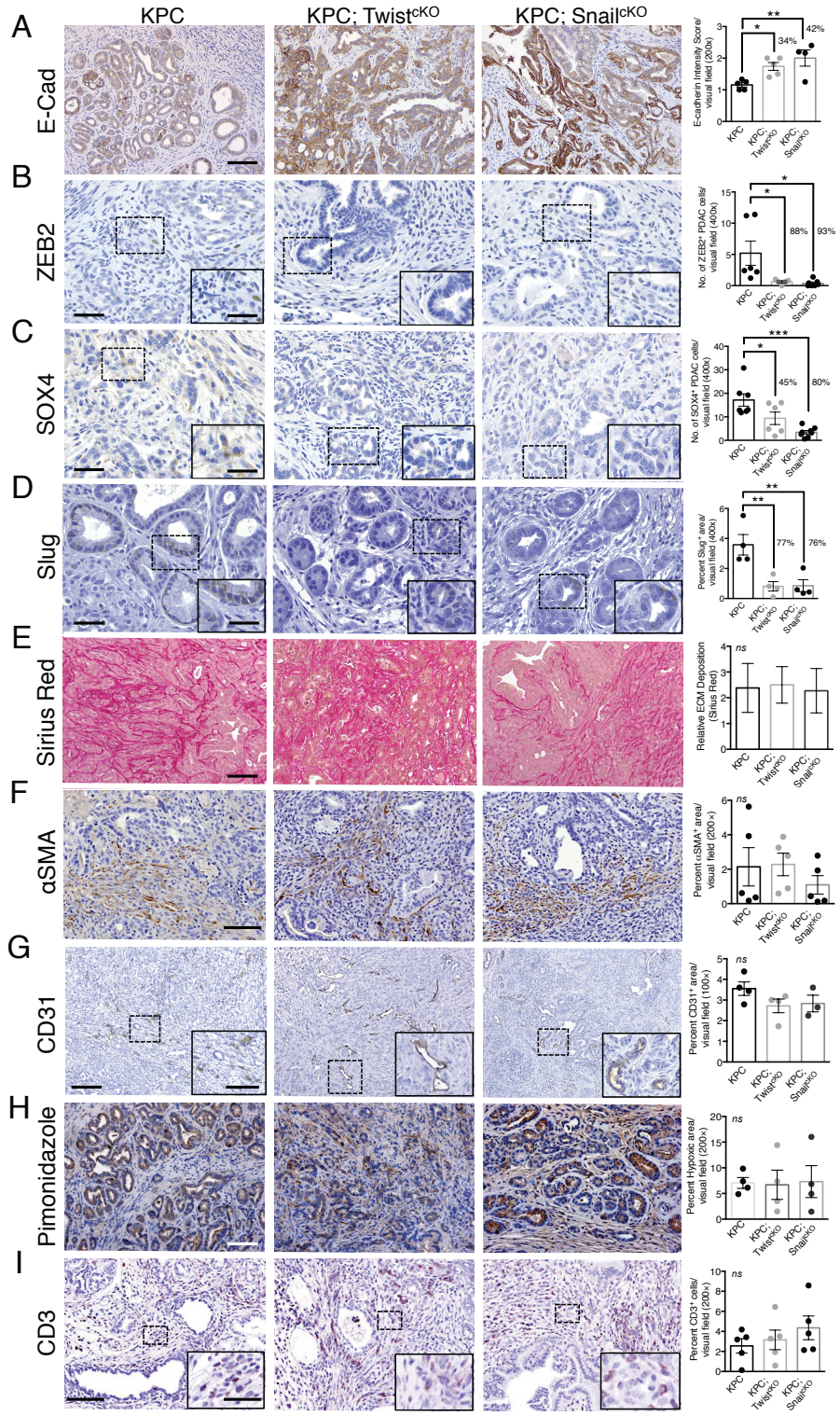
MRI analyses. MRI imaging was performed using a 7T small animal MR system as previously described[37]. To measure tumour volume, suspected regions were drawn blinded on each slice based on normalized intensities. The volume was calculated by the addition of delineated regions of interest in $mm^2 \times 1\,mm$ slice distance. None of the mice had a tumour burden that exceeded 1.5 cm in diameter, in accordance with institutional regulations. All mice with measurable tumours were enrolled in the study (see Extended Data Table 3). Mice were imaged twice, once at the beginning of the enrolment (day 0), and a second time 20 days (day 19) afterwards. Surviving animals were euthanized at end point (day 21) for histological characterization.

Statistical analyses. Statistical analyses were performed on the mean values of biological replicates in each group using unpaired two-tailed or one-tailed *t*-tests (qPCR only), or one-way ANOVA with Tukey's multiple comparisons test using GraphPad Prism, as stipulated in the figure legends. χ^2 analyses, using SPSS statistical software, were performed comparing control to cKO groups for metastatic or colonization frequency across multiple histological parameters in all mice and mice $\geq$120 days of age in Extended Data Table 1. Fisher's exact *P* value was used to determine significance. Results are outlined in Extended Data Table 2. Kaplan–Meier plots were drawn for survival analysis and the log rank Mantel-Cox test was used to evaluate statistical differences, using GraphPad Prism. Data met the assumptions of each statistical test, where variance was not equal (determined by an *F*-test) Welch's correction for unequal variances was applied. Error bars represent s.e.m. when multiple visual fields were averaged to produce a single value for each animal which was then averaged again to represent the mean bar for the group in each graph. $P < 0.05$ was considered statistically significant.

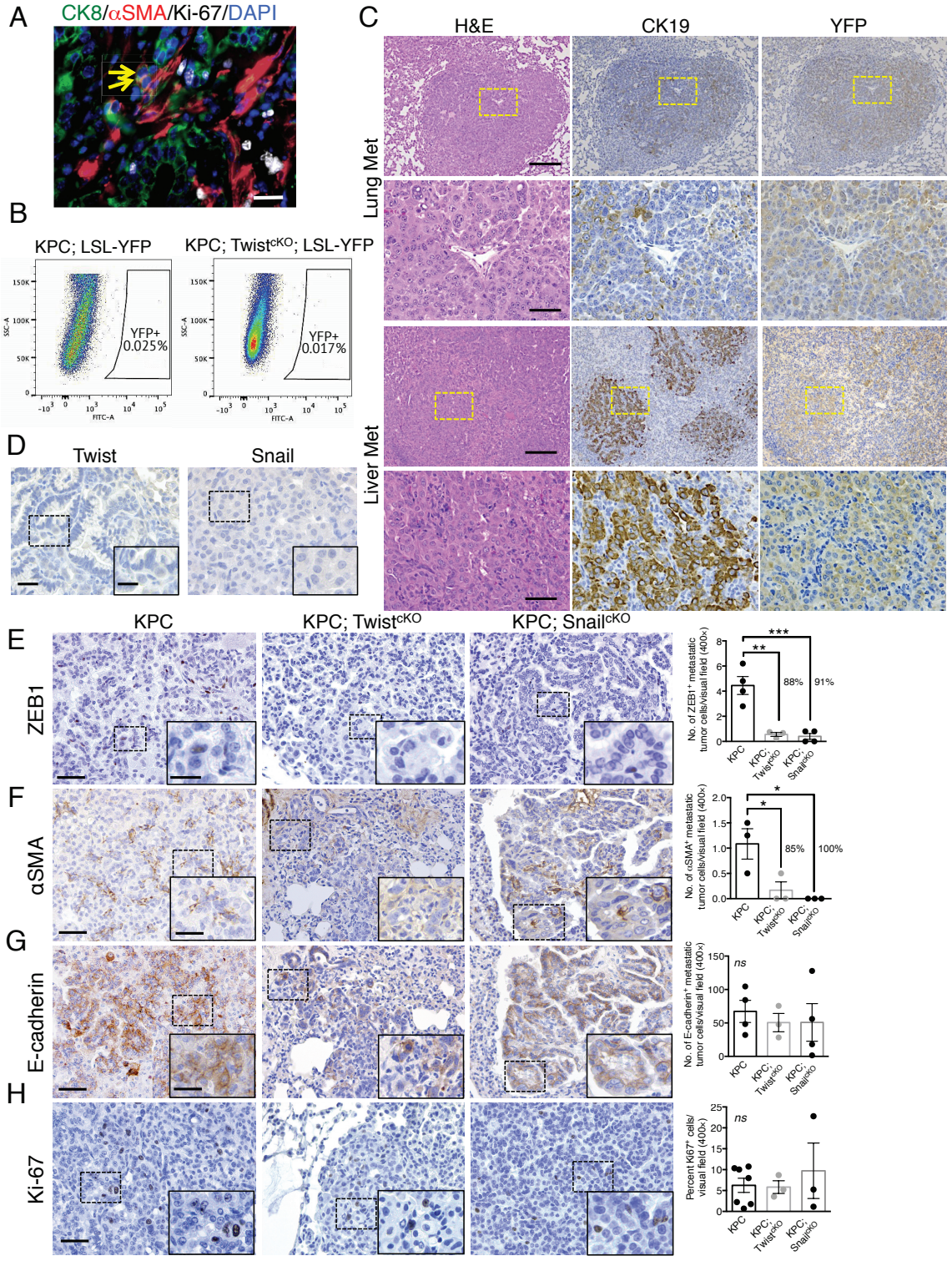
31. Hingorani, S. R. *et al.* Trp53R172H and KrasG12D cooperate to promote chromosomal instability and widely metastatic pancreatic ductal adenocarcinoma in mice. *Cancer Cell* **7,** 469–483 (2005).
32. Ijichi, H. *et al.* Aggressive pancreatic ductal adenocarcinoma in mice caused by pancreas-specific blockade of transforming growth factor-β signaling in cooperation with active Kras expression. *Genes Dev.* **20,** 3147–3160 (2006).
33. Özdemir, B. C. *et al.* Depletion of carcinoma-associated fibroblasts and fibrosis induces immunosuppression and accelerates pancreas cancer with reduced survival. *Cancer Cell* **25,** 719–734 (2014).
34. Keskin, D. *et al.* Targeting vascular pericytes in hypoxic tumors increases lung metastasis via angiopoietin-2. *Cell Rep.* **10,** 1066–1081 (2015).
35. Smyth, G. K. Limma: linear models for microarray data. In *Bioinformatics and Computational Biology Solutions using R and Bioconductor* (eds Gentleman, R., Carey, V., Dudoit, S., Irizarry, R. & Huber, W.) (Springer, 2005).
36. Ying, H. *et al.* Oncogenic Kras maintains pancreatic tumors through regulation of anabolic glucose metabolism. *Cell* **149,** 656–670 (2012).
37. Melo, S. A. *et al.* Glypican-1 identifies cancer exosomes and detects early pancreatic cancer. *Nature* **523,** 177–182 (2015).



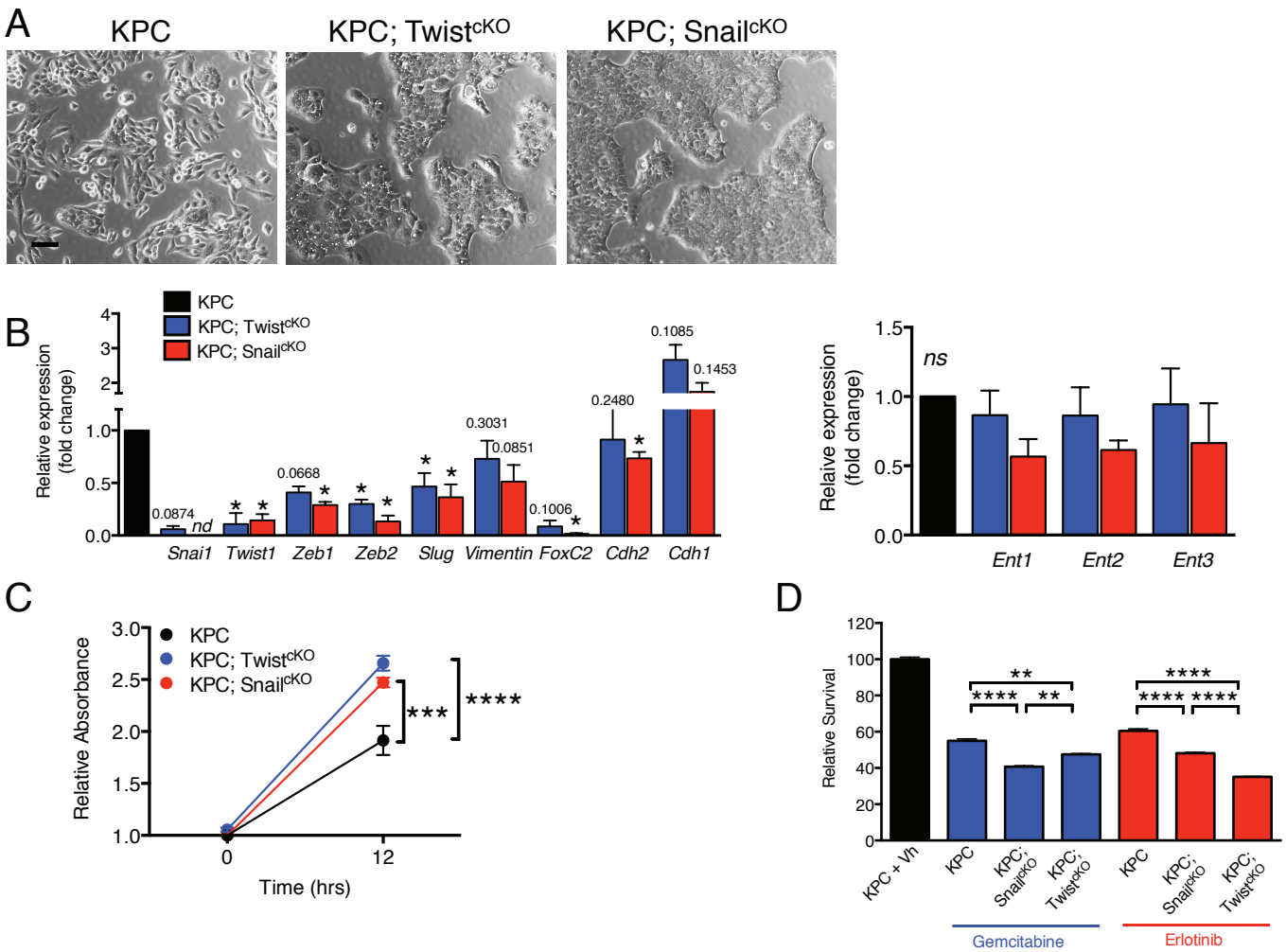
Extended Data Figure 1 | EMT inhibition is specific to tumour epithelium. a, Representative images of haematoxylin and eosin-stained small intestine (SmInt), kidney, and heart (scale bar, 100 μm). **b**, Pancreatic mass of 29 (KPC), 13 (KPC;Twist^{cKO}) and 28 (KPC;Snail^{cKO}) mice, error bars represent s.d.; one-way ANOVA. **c**, Merge of *Twist1* or *Snai1 in situ* hybridization (black) followed by CK8 (red) immunolabelling in tumours from KPC and KPC;Twist^{cKO} or KPC;Snail^{cKO} mice, respectively. White arrows highlight positive cells in the stroma, yellow arrows highlight negative epithelium (scale bar, 20 μm). **d**, Twist or Snail immunostaining in KPC and KPC;Twist^{cKO} or KPC;Snail^{cKO} tumours, respectively. Black arrows highlight positive cells in the stroma, red arrows highlight negative epithelium (scale bar, 20 μm). **e**, Channel separations of the representative images of αSMA immunolabelling in YFP lineage-traced tumours found in Fig. 1f (scale bar, 50 μm). **f**, EMT gene expression signature analysis in KPC, KPC;Twist^{cKO} and KPC;Snail^{cKO} cohorts ($n = 3$ mice). Red arrows indicate reduced *Twist1* and *Snai1* expression in KPC;Twist^{cKO} and KPC;Snail^{cKO} cohorts, respectively.



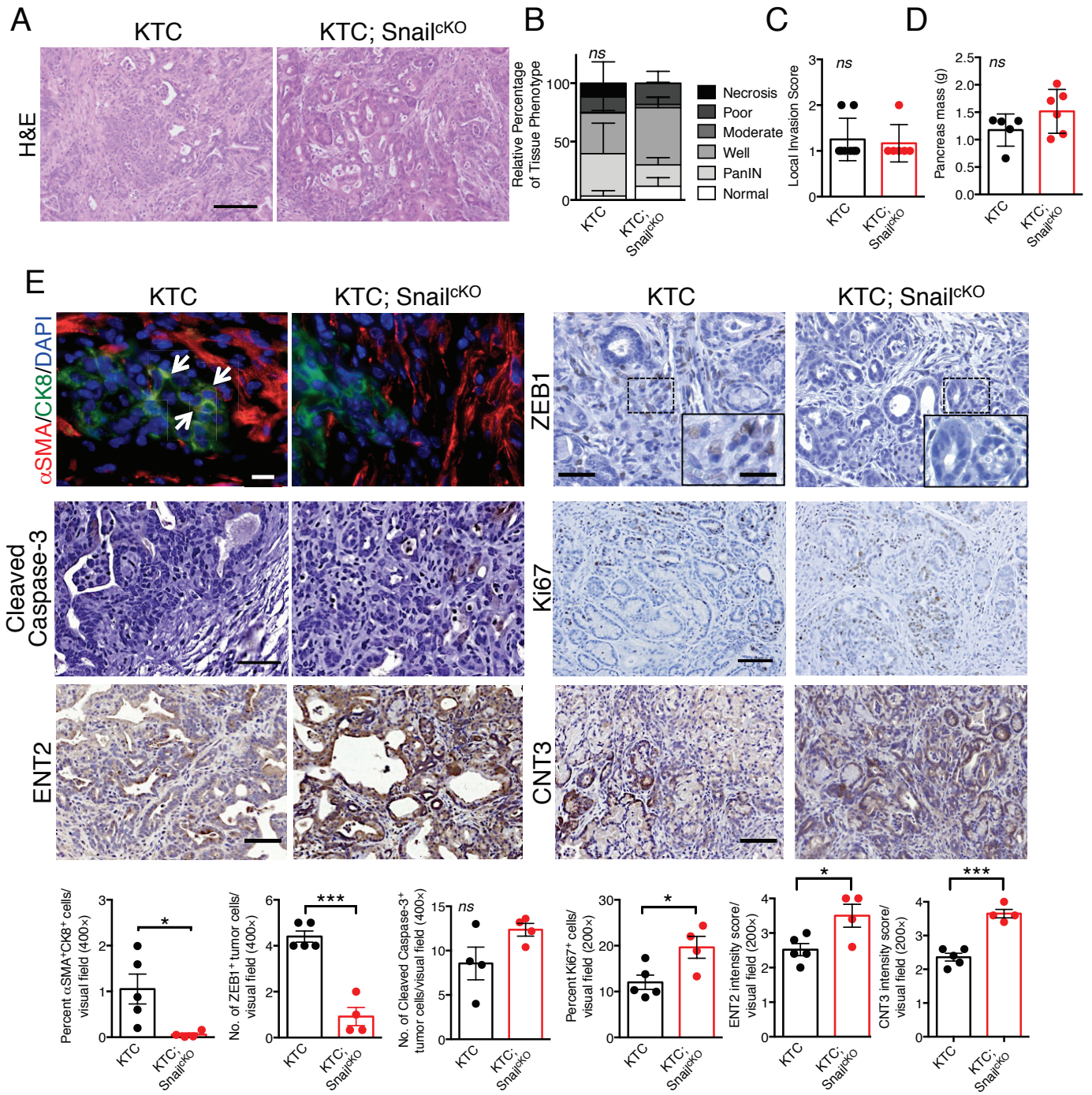
Extended Data Figure 2 | General suppression of EMT markers does not affect desmoplasia. a, E-cadherin immunolabelling and quantification of primary KPC ($n = 5$ mice), KPC;Twist^{cKO} ($n = 5$ mice) and KPC;Snail^{cKO} ($n = 4$ mice) (scale bar, 100 μm). **b**, Zeb2 immunolabelling and quantification of primary KPC ($n = 6$ mice), KPC;Twist^{cKO} ($n = 5$ mice) and KPC;Snail^{cKO} ($n = 7$ mice) (scale bar, 50 μm; inset scale bar, 20 μm). **c**, Sox4 immunolabelling and quantification of primary KPC ($n = 7$ mice), KPC;Twist^{cKO} ($n = 6$ mice) and KPC;Snail^{cKO} ($n = 8$ mice) (scale bar, 50 μm; inset scale bar, 20 μm). **d**, Slug immunolabelling and quantification of primary KPC ($n = 4$ mice), KPC;Twist^{cKO} ($n = 4$ mice) and KPC;Snail^{cKO} ($n = 4$ mice) tumours (scale bar, 50 μm; inset scale bar, 20 μm). **e**, Sirius Red staining and quantification of primary KPC ($n = 21$ mice), KPC;Twist^{cKO} ($n = 8$ mice) and KPC;Snail^{cKO} ($n = 11$ mice) (scale bar, 100 μm; error bars represent s.d.) **f**, αSMA immunolabelling and quantification of primary KPC ($n = 5$ mice), KPC;Twist^{cKO} ($n = 5$ mice) and KPC;Snail^{cKO} ($n = 5$ mice) (scale bar, 100 μm). **g**, CD31 immunolabelling and quantification of primary KPC ($n = 4$ mice), KPC;Twist^{cKO} ($n = 4$ mice) and KPC;Snail^{cKO} ($n = 3$ mice) (scale bar, 200 μm, inset scale bar, 100 μm). **h**, Pimonidazole staining and quantification of primary KPC ($n = 4$ mice), KPC;Twist^{cKO} ($n = 4$ mice) and KPC;Snail^{cKO} ($n = 4$ mice) (scale bar, 100 μm). **i**, CD3 immunolabelling and quantification of primary KPC ($n = 5$ mice), KPC;Twist^{cKO} ($n = 5$ mice) and KPC;Snail^{cKO} ($n = 5$ mice) (scale bar, 100 μm; inset scale bar, 25 μm). Unless otherwise indicated error bars represent s.e.m., and significance determined by one-way ANOVA. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$; *ns*, not significant.



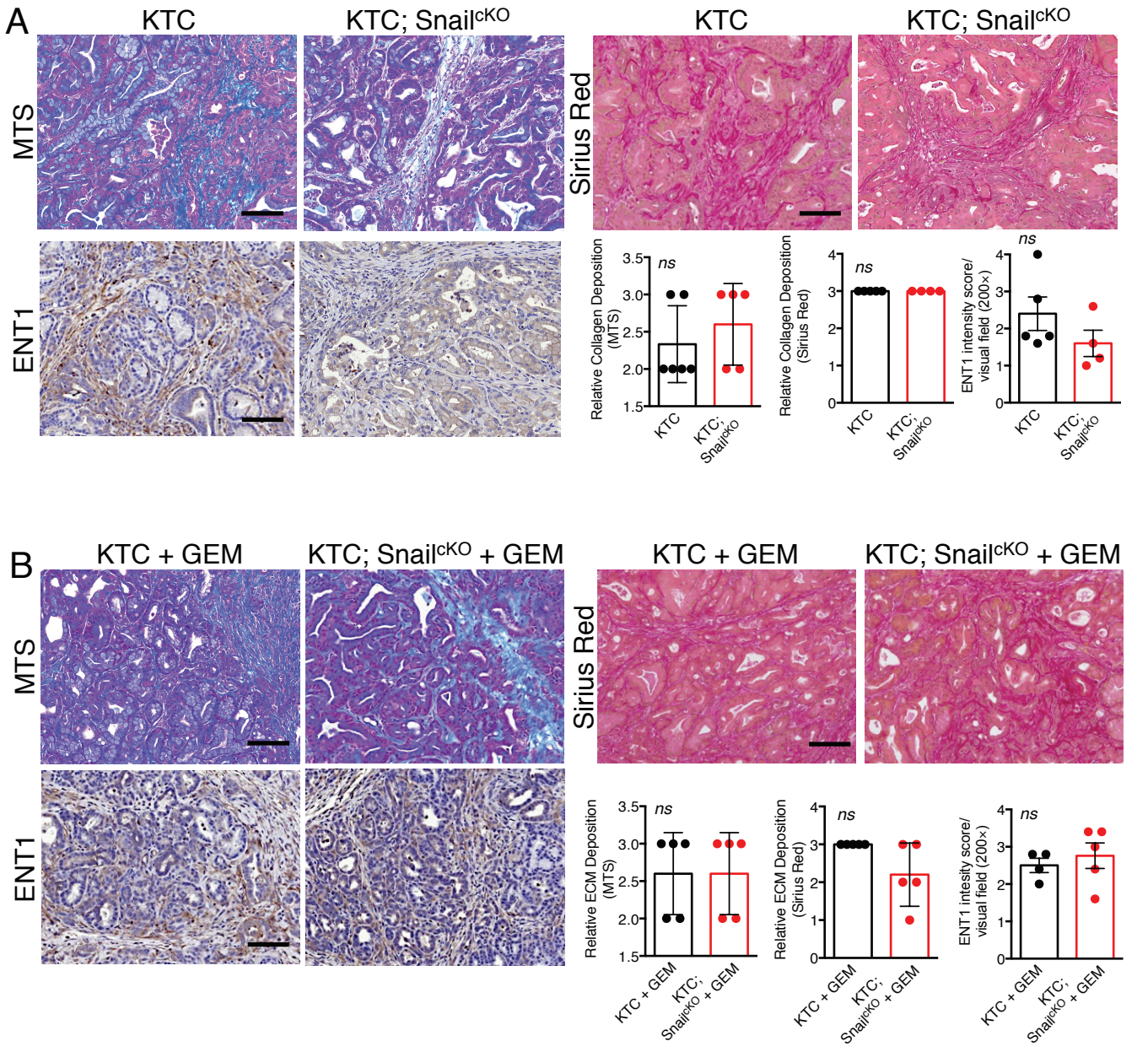
Extended Data Figure 3 | EMT suppression does not alter epithelial characteristics of metastases. a, Immunolabelling of primary tumours ($n = 3$ mice) for αSMA (red), CK8 (green), Ki67 (white) and DAPI (blue); yellow arrows indicate EMT^+ cells (scale bar, 20 μm). **b**, Representative dot plots of circulating YFP^+ cells. **c**, Images of serial sections of KPC;LSL-YFP lung and liver metastasis stained for haematoxylin and eosin or immunolabelled for CK19 or YFP. Yellow dashed box represents magnified areas in panel below (scale bar, 200 μm; magnification scale bar, 50 μm). **d**, KPC metastatic tumours stained for Twist and Snail ($n = 3$ mice; scale bar, 20 μm; inset scale bar, 10 μm). **e**, Zeb1 immunolabelling and quantification of metastatic KPC ($n = 4$ mice), KPC;$Twist^{cKO}$ ($n = 3$ mice) and KPC;$Snail^{cKO}$ ($n = 4$ mice) (scale bar, 50 μm; inset scale bar, 20 μm). **f**, αSMA immunolabelling and quantification of metastatic KPC ($n = 3$ mice), KPC;$Twist^{cKO}$ ($n = 3$ mice) and KPC;$Snail^{cKO}$ ($n = 3$ mice) (scale bar, 50 μm; inset scale bar, 20 μm). **g**, E-cadherin staining on serial sections of αSMA immunolabelling and quantification of metastatic KPC ($n = 4$ mice), KPC;$Twist^{cKO}$ ($n = 3$ mice) and KPC;$Snail^{cKO}$ ($n = 4$ mice) (scale bar, 50 μm; inset scale bar, 20 μm). **h**, Ki67 immunolabelling and quantification of metastatic KPC ($n = 7$ mice), KPC;$Twist^{cKO}$ ($n = 3$ mice) and KPC;$Snail^{cKO}$ ($n = 3$ mice) (scale bar, 50 μm; inset scale bar, 20 μm). Unless otherwise indicated error bars represent s.e.m., percentages indicated represent per cent decrease from control, and significance was determined by one-way ANOVA. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$; *ns*, not significant.



Extended Data Figure 4 | EMT suppressed primary tumour cells have reduced mesenchymal markers and show resistance to chemotherapy *in vitro*. **a**, Bright-field micrograph of cultured primary KPC, KPC;Twist^{cKO} and KPC;Snail^{cKO} cells (scale bar, 50 μm). **b**, EMT- and gemcitabine-transport-related gene expression shown by qPCR analysis in KPC ($n = 3$–4 cell lines), KPC;Twist^{cKO} ($n = 5$ cell lines) and KPC;Snail^{cKO} ($n = 5$–6 cell lines) (error bars represent s.d., one-tailed *t*-test, $^{*}P < 0.05$, numbers list non-significant *P* values. *nd*, not detected, *ns*, not significant). **c**, MTT assay showing cell proliferation in KPC, KPC;Twist^{cKO} and KPC;Snail^{cKO} cells ($n = 8$, 8 and 8 biological replicates of a cell line for each genotype). **d**, Relative cell viability (MTT assay) in cultured KPC, KPC;Twist^{cKO} and KPC;Snail^{cKO} cells treated with gemcitabine or erlotinib ($n = 8$, 8 and 8 biological replicates of a cell line for each genotype). Unless otherwise indicated error bars represent s.e.m., significance was determined by one-way ANOVA. $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$.



Extended Data Figure 5 | EMT inhibition in KTC mice mirrors phenotype observed in KPC mice. a, Representative images of haematoxylin and eosin-stained primary tumours (scale bar, 100 μm). **b**, Relative percentage of each histological tissue phenotype of KTC ($n = 8$ mice) and KTC;Snail^{cKO} ($n = 6$ mice) primary tumours (error bars represent s.d.). **c**, Primary tumour invasiveness in KTC ($n = 8$ mice) and KTC;Snail^{cKO} ($n = 6$ mice) (error bars represent s.d.). **d**, Pancreatic mass in KTC ($n = 5$ mice) and KTC;Snail^{cKO} ($n = 6$ mice) (error bars represent s.d.). **e**, Immunolabelling and quantification of primary KTC ($n = 5$ mice), KTC;Snail^{cKO} ($n = 4$ mice) for αSMA (red), CK8 (green) and DAPI (blue); white arrows indicate double-positive cells (scale bar, 20 μm), Zeb1 (scale bar, 50 μm; inset scale bar, 20 μm), cleaved caspase-3 (scale bar, 50 μm; $n = 4$ mice for both groups), Ki67 (scale bar, 100 μm), ENT2 (scale bar, 100 μm) and CNT3 (scale bar, 100 μm); error bars represent s.e.m. Significance was determined by two-tailed t-test. $^*P < 0.05$, $^{***}P < 0.001$; *ns*, not significant.



Extended Data Figure 6 | Desmoplasia is unaffected in EMT suppressed tumours with or without gemcitabine. a, **b**, Staining and quantification of KTC ($n = 5$ or 6 mice), KTC;Snail^{cKO} ($n = 4$ or 5 mice), KTC plus gemcitabine (+ GEM; $n = 4$ or 5 mice), KTC;Snail^{cKO} + GEM ($n = 5$ mice) for Masson's trichrome stain (MTS) (scale bars, 100 μm), Sirius Red staining (scale bars, 100 μm), and ENT1 (scale bars, 100 μm). Error bars represent s.d. (MTS and Sirius Red) or s.e.m. (ENT1), and significance was determined by two-tailed *t*-test. *ns*, not significant.

Extended Data Table 1 | Pathological spectrum of primary disease and metastasis in KPC, KPC;TwistcKO and KPC;Snail^{cKO} cohorts

Pathological Spectrum within cohorts

ID	AGE	PDA	Differentiation	Histology 1	Histology 2	Liver	Lung	Spleen	Any	Moribund
KPC	**(104)**									
1	158	Y	W	S	G	Y	Y	N	Y	Y
2	165	Y	W	G		N	N	N	N	Y
3	148	Y	P	S	G	N	N	-	N	Y
4	135	Y	M	S	G	Y	N	Y	Y	Y
5	95	Y	M	G		N	Y	N	Y	N
6	42	Y	M	G		N	N	N	N	Y
7	55	Y	P	G	S	Y	N	N	Y	Y
8	91	Y	M	G		N	N	N	N	N
9	87	Y	W	G		N	N	N	N	N
10	63	Y	P	G		Y	Y	Y	Y	N
11	108	Y	P	S	G	Y	N	N	Y	FD
12	110	Y	W	G		N	N	N	N	N
13	104	Y	W	G		Y	N	N	Y	Y
14	54	Y	W	S	G	N	N	N	N	Y
15	108	Y	P	S	G	N	Y	N	Y	Y
16	42	Y	P	S	G	N	N	N	N	Y
17	68	Y	W	G		N	N	N	N	N
18	107	Y	P	G		N	N	N	N	N
19	87	Y	P	G		N	N	N	N	N
20	48	Y	P	G	S	N	N	N	N	Y
21	109	Y	P	G	S	Y	Y	N	Y	FD
22	81	Y	P	G		Y	Y	N	Y	Y
23	151	Y	W	G		N	Y	N	Y	Y
24	47	Y	M	G	S	N	Y	N	Y	Y
25	143	Y	P	G	S	N	Y	N	Y	Y
26	122	Y	W	G		Y	N	N	Y	N
27	115	Y	P	G		Y	Y	N	Y	N
28	76	Y	W	G		N	Y	N	Y	N
29	122	Y	M	S	G	Y	N	N	Y	Y
30	97	Y	P	G		N	N	N	N	N
31	107	Y	W	G		N	N	N	N	N
Totals	(Median)	31/31				11/31	11/31	2/30	17/31	
%		100.0%				35.5%	35.5%	6.7%	54.8%	
Twist^{cKO}	**(111)**									
1	148	Y	W	G	S	Y	N	N	Y	N
2	151	Y	P	S	G	Y	Y	Y	Y	N
3	140	Y	P	G		Y	Y	N	Y	Y
4	53	Y	P	G	S	N	N	N	N	Y
5	43	Y	P	G		N	N	N	N	Y
6	117	Y	P	G	S	N	N	N	N	N
7	90	Y	P	S	G	Y	N	N	Y	Y
8	52	Y	P	G	S	N	N	N	N	Y
9	104	Y	P	G		N	N	N	N	N
10	218	Y	P	G	S	N	N	Y	Y	Y
11	153	Y	P	G		N	Y	N	Y	Y
12	45	Y	P	G	S	N	N	N	N	Y
13	77	Y	P	G	S	Y	N	N	Y	Y
14	126	Y	P	G	S	Y	Y	N	Y	Y
Totals	(Median)	14/14				6/14	4/14	2/14	8/14	
%		100.0%				42.9%	28.6%	14.3%	57.1%	
Snail^{cKO}	**(103)**									
1	144	Y	W	G		N	Y	N	Y	N
2	51	Y	P	G	S	N	N	N	N	Y
3	105	Y	P	G	S	N	Y	N	Y	Y
4	111	Y	P	G		N	N	N	N	N
5	106	Y	P	G	S	Y	N	Y	Y	Y
6	129	Y	P	G		N	N	N	N	N
7	102	Y	P	G	S	N	Y	-	Y	N
8	98	Y	P	G	S	Y	N	Y	Y	N
9	47	Y	P	G	S	N	N	N	N	Y
10	54	Y	W	G		Y	Y	N	Y	FD
11	59	Y	M	G		Y	N	N	Y	N
12	103	Y	P	G		Y	N	N	Y	N
13	60	Y	P	S	G	Y	N	Y	Y	Y
14	77	Y	P	G		Y	N	N	Y	Y
15	57	Y	M	S	G	Y	N	N	Y	FD
16	130	Y	P	G		Y	Y	N	Y	FD
17	76	Y	P	G	S	N	N	N	N	FD
18	111	Y	P	G		N	Y	N	Y	Y
19	100	Y	P	G	S	Y	N	Y	Y	FD
20	104	Y	P	G	S	Y	N	N	Y	Y
21	124	Y	M	G		N	N	N	N	FD
22	88	Y	P	G	S	N	N	N	N	Y
23	192	Y	W	G		Y	Y	N	Y	Y
24	122	Y	P	G		N	N	N	N	Y
25	60	Y	W	G	S	N	N	N	N	Y
26	112	Y	W	G		N	Y	N	Y	N
27	48	Y	P	G	S	N	N	N	N	Y
28	48	Y	P	G	S	N	N	N	N	Y
29	124	Y	P	G	S	Y	Y	Y	Y	N
30	215	Y	W	G		N	N	N	N	N
Totals	(Median)	30/30				13/30	9/30	5/29	18/30	
%		100.0%				43.3%	30.0%	17.2%	60.0%	

Y, yes; N, no; W, well; M, moderate; P, poor; G, glandular; S, sarcomatoid; FD, found dead; –, no tissue.

Extended Data Table 2 | Results of χ^2 analysis of KPC cohorts in Extended Data Table 1

χ2 Analysis		
Group	**Parameter**	**Fisher's Exact *P* value**
Differentiation	**All Ages**	
Control vs. Twist^{cKO}	Early Tumor progression	0.458
Control vs. Snail^{cKO}		0.106
Control vs. Twist^{cKO}	Late Tumor progression	0.458
Control vs. Snail^{cKO}		0.106
Control vs. Twist^{cKO}	Sarcomatoid	0.108
Control vs. Snail^{cKO}		0.446
Differentiation	**≥ 120 days**	
Control vs. Twist^{cKO}	Early Tumor progression	0.580
Control vs. Snail^{cKO}		0.569
Control vs. Twist^{cKO}	Late Tumor progression	0.580
Control vs. Snail^{cKO}		0.569
Control vs. Twist^{cKO}	Sarcomatoid	1.000
Control vs. Snail^{cKO}		0.119
Metastasis	**All Ages**	
Control vs. Twist^{cKO}	Liver Metastasis	0.744
Control vs. Snail^{cKO}		0.605
Control vs. Twist^{cKO}	Lung Metastasis	0.743
Control vs. Snail^{cKO}		0.786
Control vs. Twist^{cKO}	Spleen Invasion	0.581
Control vs. Snail^{cKO}		0.254
Control vs. Twist^{cKO}	Any Metastasis	1.000
Control vs. Snail^{cKO}		0.797
Metastasis	**≥ 120 days**	
Control vs. Twist^{cKO}	Liver Metastasis	0.627
Control vs. SnailcKO		1.000
Control vs. Twist^{cKO}	Lung Metastasis	0.592
Control vs. SnailcKO		1.000
Control vs. Twist^{cKO}	Spleen Invasion	0.559
Control vs. SnailcKO		1.000
Control vs. Twist^{cKO}	Any Metastasis	0.473
Control vs. SnailcKO		0.608

Extended Data Table 3 | Survival and primary tumour burden determined by MRI in KPC, KPC;Twist^{cKO} and KPC;Snail^{cKO} cohorts treated with gemcitabine

KPC Gemcitabine cohorts

ID	Start Age (Days)	Start Volume (mm^3)	End Volume (mm^3)	Survival (Days)
KPC + GEM	**(89)**			**(13)**
1	148	1610.4	D	7
2	72	29.7	D	13
3	72	439.8	902.8	21*
4	80	44.1	D	14
5	100	536.3	592.3	21*
6	89	167.0	D	2
7	94	52.7	D	7
8	122	90.2	D	14
9	164	217.9	D	8
10	143	212.8	D	18
11	84	323.8	897.2	21*
12	58	76.7	D	4
13	58	116.2	D	8
Mean	(Median)	301.4	797.4	
Stdev		406.9	145.1	
Twist^{cKO} + GEM	**(79)**			**(21)**
1	117	243.0	644.2	21*
2	75	47.2	180.0	21*
3	75	45.4	460.9	21*
4	78	54.6	47.5	21*
5	46	53.7	66.5	21
6	96	63.1	D	13
7	90	23.9	D	13
8	79	101.0	D	14
9	52	28.5	D	14
10	52	49.4	98.706	21*
11	104	43.4	127.0	21*
12	104	53.5	12.1	21*
13	68	56.7	D	15
14	122	650.1	164.1	21*
15	104	181.8	78.6	21*
Mean	(Median)	113.0	187.9	
Stdev		154.8	193.0	
Snail^{cKO} + GEM	**(96)**			**(21)**
1	188	255.2	D	12
2	181	854.7	D	4
3	127	32.0	59.6	21*
4	127	58.7	107.4	21*
5	142	109.8	D	14
6	54	33.6	57.2	21*
7	89	17.0	D	13
8	78	54.9	39.6	21*
9	78	3.1	D	15
10	104	209.7	134.3	21*
11	96	220.0	280.2	21*
12	96	24.1	46.2	21*
13	119	711.0	D	18
14	126	655.6	805.4	21*
15	119	168.6	D	18
16	82	453.8	517.4	21*
17	82	56.7	74.1	21*
18	90	40.0	D	16
19	67	80.5	D	10
20	66	49.5	226.2	21*
Mean	(Median)	204.4	213.4	
Stdev		250.7	231.7	

D, died; *euthanized at end point.

LETTER

doi:10.1038/nature15767

In situ structures of the segmented genome and RNA polymerase complex inside a dsRNA virus

Xing Zhang[1]*, Ke Ding[2,3]*, Xuekui Yu[2]*, Winston Chang[1], Jingchen Sun[2,4] & Z. Hong Zhou[1,2,3]

Viruses in the *Reoviridae*, like the triple-shelled human rotavirus and the single-shelled insect cytoplasmic polyhedrosis virus (CPV), all package a genome of segmented double-stranded RNAs (dsRNAs) inside the viral capsid and carry out endogenous messenger RNA synthesis through a transcriptional enzyme complex (TEC)[1]. By direct electron-counting cryoelectron microscopy and asymmetric reconstruction, we have determined the organization of the dsRNA genome inside quiescent CPV (q-CPV) and the *in situ* atomic structures of TEC within CPV in both quiescent and transcribing (t-CPV) states. We show that the ten segmented dsRNAs in CPV are organized with ten TECs in a specific, non-symmetric manner, with each dsRNA segment attached directly to a TEC. The TEC consists of two extensively interacting subunits: an RNA-dependent RNA polymerase (RdRP) and an NTPase VP4. We find that the bracelet domain of RdRP undergoes marked conformational change when q-CPV is converted to t-CPV, leading to formation of the RNA template entry channel and access to the polymerase active site. An amino-terminal helix from each of two subunits of the capsid shell protein (CSP) interacts with VP4 and RdRP. These findings establish the link between sensing of environmental cues by the external proteins and activation of endogenous RNA transcription by the TEC inside the virus.

Each capsid of viruses in the *Reoviridae* contains 9–12 segmented dsRNAs and up to 12 TECs. These RNA-containing viruses are fully capable of RNA transcribing and capping[1]. Crystal structures of the RdRP component of the TEC have been determined for rotavirus and mammalian reovirus (MRV)[2,3], but no high-resolution *in situ* structure of the TEC is available. Moreover, the organization of TECs with the dsRNA genome and the mechanism of transcriptional activation have remained unresolved, in contrast to the well understood genome organization inside dsDNA viruses[4,5].

With only a single protein shell that encloses ten different genome segments, CPV is one of the simplest dsRNA viruses[6] and serves as a model system, as highlighted by its contribution to the discovery of RNA capping[7]. To gain insight into the organization of the TEC and segmented dsRNA genome, we have determined CPV structure in a quiescent (q-CPV) state at 5.1 Å resolution (see Methods and Extended Data Figs 1 and 2). The structure reveals that each CPV contains ten TECs under ten specific positions of the twelve icosahedral vertices (Fig. 1). The two vertices without TECs are occupied by rod-like densities (Fig. 1a–e, Supplementary Video 1 and Extended Data Figs 3 and 4). The previously ambiguous locations of TECs[8,9] are now determined to be ten specific positions in each CPV particle, related by incomplete-D3 symmetry, with only one on a 'south tropic' position and three each around the 'north tropic', 'north pole' and 'south pole' positions (Fig. 1d and Supplementary Video 2).

Each TEC is surrounded by rod-like densities with lengths up to ~650 Å (Fig. 1a–c, e–f, Extended Data Fig. 4a and Supplementary Video 1). In most regions, these rods form parallel striations with an inter-rod distance of ~27 Å, as suggested previously (for example, refs 10, 11). Some of the rods exhibit the characteristic minor and major grooves typical of dsRNA duplex (Fig. 1g). We therefore interpret these rod-like densities as dsRNA duplexes. Unlike the model of each genome segment spiralling around one TEC (ref. 12), the duplexes do not spiral locally

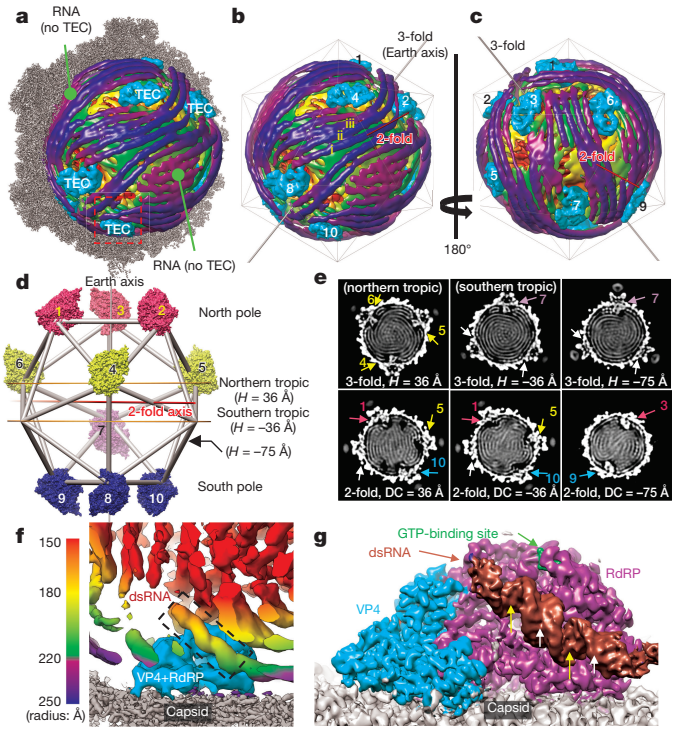


Figure 1 | Transcription enzyme complex and dsRNA genome organization inside CPV. a, Superposition of the high-resolution (3.9 Å) map of half a capsid (grey) and low-resolution (22 Å) map of dsRNA genome (radially coloured as in **f**) and TECs (cyan). **b**, **c**, Front (**b**) and back (**c**) views of the dsRNA genome and TECs of panel **a**. **d**, Earth-like representation, illustrating the locations of the ten TECs (surface-rendered) with pseudo-D3 symmetry: three on each pole and the northern tropic but only one on the southern tropic. **e**, Cross-sections of the 22 Å density map, perpendicular to either the 'earth axis' in **d** (top row) or a D3 two-fold axis (bottom row). Densities of TECs are numbered as in **d**, and the two vertices without TEC but with RNA are indicated by white arrows. DC, distance from centre. **f**, Boxed region in **a** containing RNA threads (radially coloured as in the bar) and a TEC (cyan) with bound dsRNA (dashed box). **g**, Averaged TEC region, filtered to 4.5 Å and viewed as the southern-most TEC of **a**. The RdRP-bound dsRNA has the same structure in all TECs and shows major (yellow arrows) and minor (white arrows) grooves.

[1]California Nanosystems Institute, University of California, Los Angeles, California 90095, USA. [2]Department of Microbiology, Immunology and Molecular Genetics, University of California, Los Angeles, California 90095, USA. [3]Bioengineering, University of California, Los Angeles, California 90095, USA. [4]Subtropical Sericulture and Mulberry Resources Protection and Safety Engineering Research Center, Guangdong Provincial Key Laboratory of Agro-animal Genomics and Molecular Breeding, College of Animal Science, South China Agricultural University, Guangzhou, Guangdong 510642, China.
*These authors contributed equally to this work.

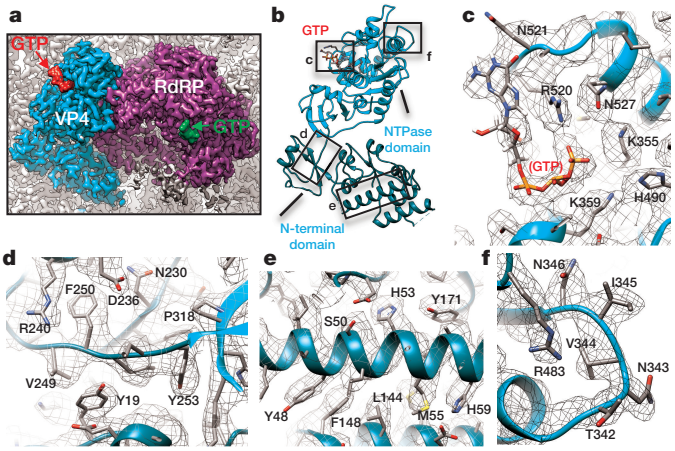


Figure 2 | Averaged TEC map at 3.3 Å resolution and *de novo* modelling of VP4. a, Averaged map of the TEC region showing VP4 (cyan) and RdRP (purple), both anchored to the inner surface of the capsid (grey). **b**, Atomic model of VP4. **c**–**f**, The boxed regions in **b**, showing density (meshes) superposed with atomic models of the GTP-binding site (**c**), a loop (**d**), a helix (**e**) and an RdRP-interacting loop (**f**).

around TECs (Fig. 1a–c, Extended Data Fig. 4 and Supplementary Video 1); instead, many extend tangentially from one TEC to another (for example, duplexes i–iii in Fig. 1b; see also Extended Data Fig. 4), indicating that each dsRNA segment is organized beyond one TEC. Indeed, the whole RNA genome is organized into seven to eight non-concentric layers with visible connections between adjacent layers (Fig. 1e and Extended Data Fig. 3). This extended organization of dsRNA is consistent with the rather long (~620 Å) persistence length of dsRNA[13] and would reduce the energy needed for genome packaging and transcription. One RNA duplex (the brown one in Fig. 1g) binds to each of the ten TECs at the same relative position and orientation, suggesting that this RNA duplex is a conserved feature among the ten dsRNA segments. However, the organization of the remaining RNA duplex differs among the ten TECs (Extended Data Fig. 4g). The two vertices without TECs are occupied only by roughly parallel dsRNA densities (Fig. 1a–c).

We also obtained a 3.9 Å resolution asymmetric reconstruction directly from the raw images of q-CPV and subsequently used non-crystallographic averaging to improve the resolution to 3.3 Å for the TEC-containing regions (see Methods and Extended Data Fig. 5). The averaged map retains a short (~35 Å) RdRP-bound dsRNA density (Fig. 1g) and resolves the two protein components of the TEC: VP4 and RdRP (Fig. 2a). We built a backbone model of the RdRP-bound dsRNA and *de novo* atomic models of both VP4 and RdRP (Fig. 2c–f, Extended Data Fig. 6 and Supplementary Videos 3–8). VP4 and RdRP interact extensively (Fig. 2a) with a buried interface area of ~2,800 Å^2.

VP4 appears 'L'-shaped and consists of an amino-terminal (amino acids 1–252) and a carboxy-terminal (253–561) domain, with two unresolved/flexible segments (amino acids 23–40 and 86–131) (Fig. 2a, b and Supplementary Video 3). The N-terminal domain is formed by two small β-sheets and several α-helices, and the main body of the C-terminal domain is a Walker-A α/β motif, a well-known NTP-binding motif found in the P-loop kinase family of proteins. Sequence analysis predicted an NTP binding site in VP4 (refs 14, 15). Indeed, the VP4 structure contains a GTP molecule at the predicted NTP binding site of the C-terminal domain (Fig. 2c and Supplementary Video 4). We thus rename the C-terminal domain as the NTPase domain (Fig. 2b, c). A similar fold was also observed in the N-terminal α/β domain of bluetongue virus VP4. But, remarkably, bluetongue virus VP4 is an RNA capping enzyme and its α/β domain does not bind GTP (ref. 16). CPV VP4 and its homologues in other dsRNA viruses have been speculated to function as an NTPase, as an RNA 5′-triphosphatase (RTPase) or as a helicase[14,17,18]. Our structure supports VP4 as an NTPase but shows no interaction with dsRNA, suggesting that VP4 is unlikely to be a helicase. Whether VP4 is the CPV RTPase or an RdRP regulatory factor remains to be determined.

Like other RdRP structures[2,3,19], the CPV RdRP contains a polymerase core with finger (amino acids 349–515, 549–641), thumb (730–863) and palm (516–548, 642–729) subdomains (Fig. 3a). This polymerase core is sandwiched between the N-terminal (1–348) and C-terminal bracelet (864–1225) domains (Fig. 3 and Extended Data Fig. 6). A GTP is identified (Figs 1g and 3a) at the position equivalent to the cap-binding site observed in the MRV RdRP (ref. 2). Interestingly, the bracelet domain of q-CPV RdRP differs from that of MRV significantly, despite close similarities between both their polymerase core and their N-terminal domains. Consequently, the crystal structure of MRV RdRP has an open RNA template entry channel and an accessible polymerase active site[2]; while in the q-CPV RdRP, the polymerase active site is covered by the bracelet domain and there is no recognizable channel for template entry (Figs 3a and 4a and Supplementary Videos 5 and 8). Since q-CPV is incapable of mRNA transcription, we considered that these structural differences might be characteristic of conformational differences between bracelet-containing RdRPs in the quiescent and transcribing states.

To test this hypothesis, we then determined the structure of actively transcribing CPV (t-CPV), obtained an averaged TEC map at 4.0 Å resolution, and built atomic models of VP4 and RdRP (Fig. 3b, Extended Data Figs 5 and Supplementary Video 9). In t-CPV, the location of TECs remains the same, as do the structures of VP4 and those of the N-terminal and polymerase core domains of RdRP (Fig. 3a–f, Extended Data Figs 7–9 and Supplementary Video 10). By contrast, the RdRP bracelet domain undergoes major conformational change (Fig. 3d, e). Consistent with the above hypothesis, the *in situ* structure of the t-CPV RdRP is quite similar to the crystal structure of MRV RdRP in its elongation state[2] (Extended Data Fig. 10).

The most significant changes of the CPV RdRP between quiescent and transcribing states involve two neighbouring structural modules in the bracelet domain: the capsid-proximal module A (amino acids 1080–1140 containing helices Bα14–Bα16) and the

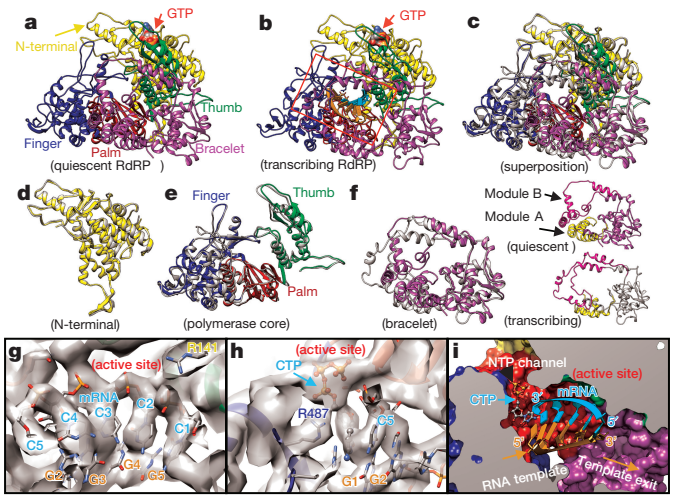


Figure 3 | Comparison of RdRP in quiescent and transcribing states. **a**, **b**, Ribbon models of RdRP in quiescent (**a**) and transcribing (**b**) states. The latter contains fragments of RNA template (orange) and nascent mRNA (cyan) inside the active site (box). **c**–**f**, Superpositions of RdRP structures in quiescent (colour) and transcribing (grey) states shown in full (**c**) and as separate domains—N-terminal (**d**), polymerase (**e**), and bracelet (**f**) with modules A (yellow) and B (magenta) further highlighted on its right panel. **g**–**i**, Densities (grey) and models (ribbons and sticks) of nucleic acids in the active site of t-CPV RdRP. The fragments of the (−)RNA template and the nascent mRNA in the active site are modelled as a poly-G and poly-C, respectively. In **h**, a CTP is placed in the NTP-binding site and in **i**, the template and mRNA form RNA duplex in the active site of RdRP (surface-rendered model).

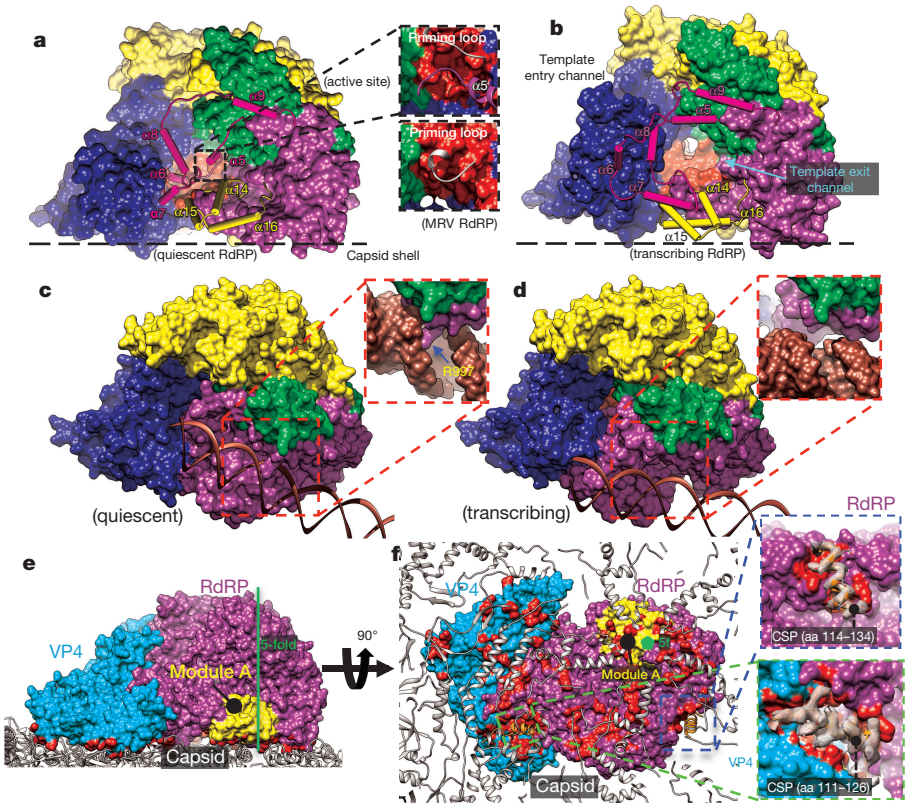


Figure 4 | Interactions between TEC and CSPs. a, **b**, Conformational changes of modules A (yellow loops/helices as wires/cylinders) and B (magenta loops/helices as wires/cylinders) in quiescent (**a**) and transcribing (**b**) states. Module A interacts with the capsid shell, and the loop-Bα5 fragment of module B blocks the active site (inset) in the quiescent state (**a**) but retracts to expose the active site in the transcribing (**b**) state (see Extended Data Fig. 7). **c**, **d**, The RdRP-bound dsRNA (ribbon) in the quiescent state (**c**) is detached from RdRP in the transcribing state (**d**). **e**, **f**, Interactions of CSPs (ribbons) with RdRP (purple and yellow) and VP4 (cyan). Residues of RdRP and VP4 within 4 Å distance to the capsid shell are marked in red. An icosahedral five-fold axis is indicated by a green line in **e** and a green pentagon in **f**. Insets in **f** indicate two CSP N-terminal helices (white density with ribbon-and-stick models): one (upper) interacts only with RdRP while the other (lower) with both RdRP and VP4.

VP4-proximal module B (912–1010 containing helices Bα5–Bα9) (Fig. 4a, b and Extended Data Figs 9 and 10). Compared to that in q-CPV, module A in t-CPV rotates ~40° towards the capsid shell (Fig. 3f and Extended Data Figs 9 and 10f–k). Consistent with previous icosahedral reconstructions, our asymmetric reconstructions show that the capsid shell of t-CPV expands outwards from q-CPV, with the maximal (~10 Å) expansion occurring at the vertex region[20,21], to which module A of the bracelet domain is attached (Fig. 4e, f). Likewise, module B refolds substantially from quiescent to transcribing state, such that a template entry channel is formed (Fig. 4a, b) and the blockage of the active site by the Bα5-loop-Bα6 fragment is removed (Figs 3f, 4a, b and Extended Data Figs 9 and 10).

In the quiescent state, a helical dsRNA duplex is held inside a shallow cleft formed by modules A and B (Figs 1g, 4c and Extended Data Fig. 7d, f) through interaction between a major groove of the RNA duplex and residue Arg997 of module B (Fig. 4c, inset). In the transcribing state, this RNA duplex becomes detached, perhaps as a result of refolding the RdRP bracelet domain (Fig. 4d and Extended Data Fig. 7e, g). We reason that detachment of the RNA duplex would permit RNA to slide towards the template entry channel for RNA synthesis in t-CPV. Indeed, in the catalytic centre of the t-CPV RdRP, we observe weak densities (Fig. 3b, g–i) that match the RNA duplex in the crystal structure of the MRV RdRP elongation complex[2]. We are able to place a 5-base-pair (bp) RNA backbone model in the active site and a CTP at the NTP binding site (Fig. 3g–i).

In addition to enclosing the viral genome and anchoring TECs, the CSP also regulates polymerase activity in dsRNA viruses[22–25]. In particular, the CSP N-terminal fragment is involved in genome replication, mRNA transcription and capping[23,26,27]. A CSP N-terminal fragment, unresolved in all previous structures[21,28–30], is resolved here to form a helix in the two TEC-interacting CSP subunits in both q-CPV and t-CPV (Fig. 4e, f). The N-terminal helix of one CSP inserts into the interface between the NTPase domain of VP4 and the finger subdomain of RdRP (Fig. 4f, lower inset), and that of the other CSP interacts with the bracelet domain of RdRP (Fig. 4f, upper inset). Notably, the former is in proximity to the NTP-binding site of the VP4 NTPase, suggesting how the N-terminal fragment of CSP is positioned to affect TEC. In addition, the structures reveal that other regions (that is, areas under the vertex) of CSP also interact with module A of the RdRP bracelet domain (Fig. 4e, f). From quiescent to transcribing state, module A and the CSP regions involved in this interaction both undergo conformational changes. Taken together, these results point to a sequence of conformational changes that leads to activation of endogenous transcription. Specifically, environmental cues cause the capsid shell to expand[21], which triggers refolding of the RdRP bracelet domain, leading to formation of the entry channel for a RNA template and exposure of the polymerase active site for RNA synthesis.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 8 June; accepted 7 October 2015.
Published online 26 October 2015; corrected online 25 November 2015 (see full-text HTML version for details).

1. Mertens, P. P. C., Rao, S. & Zhou, Z. H. in *Virus Taxonomy, VIIIth Report of the ICTV* (eds Fauquet, C. M. *et al.*) 522–533 (Elsevier/Academic Press, 2004).
2. Tao, Y., Farsetta, D. L., Nibert, M. L. & Harrison, S. C. RNA synthesis in a cage—structural studies of reovirus polymerase λ3. *Cell* **111,** 733–745 (2002).
3. Lu, X. *et al.* Mechanism for coordinated RNA packaging and genome replication by rotavirus polymerase VP1. *Structure* **16,** 1678–1688 (2008).
4. Jiang, W. *et al.* Structure of epsilon15 bacteriophage reveals genome organization and DNA packaging/injection apparatus. *Nature* **439,** 612–616 (2006).

5. Lander, G. C. *et al.* The structure of an infectious P22 virion shows the signal for headful DNA packaging. *Science* **312,** 1791–1795 (2006).
6. Zhou, Z. H. in *Segmented Double-Stranded RNA Viruses: Structure and Molecular Biology* (ed. Patton, J. T.) 27–43 (Caister Academic Press, 2008).
7. Furuichi, Y. "Methylation-coupled" transcription by virus-associated transcriptase of cytoplasmic polyhedrosis virus containing double-stranded RNA. *Nucleic Acids Res.* **1,** 809–822 (1974).
8. Estrozi, L. F. *et al.* Location of the dsRNA-dependent polymerase, VP1, in rotavirus particles. *J. Mol. Biol.* **425,** 124–132 (2013).
9. Zhang, X., Walker, S. B., Chipman, P. R., Nibert, M. L. & Baker, T. S. Reovirus polymerase λ3 localized by cryo-electron microscopy of virions at a resolution of 7.6Å. *Nature Struct. Mol. Biol.* **10,** 1011–1018 (2003).
10. Nason, E. L. *et al.* Interactions between the inner and outer capsids of bluetongue virus. *J. Virol.* **78,** 8059–8067 (2004).
11. Xia, Q., Jakana, J., Zhang, J.-Q. & Zhou, Z. H. Structural comparisons of empty and full cytoplasmic polyhedrosis virus protein-RNA interactions and implications for endogenous RNA transcription mechanism. *J. Biol. Chem.* **278,** 1094–1100 (2003).
12. Gouet, P. *et al.* The highly ordered double-stranded RNA genome of bluetongue virus revealed by crystallography. *Cell* **97,** 481–490 (1999).
13. Abels, J., Moreno-Herrero, F., Van der Heijden, T., Dekker, C. & Dekker, N. Single-molecule measurements of the persistence length of double-stranded RNA. *Biophys. J.* **88,** 2737–2744 (2005).
14. Nibert, M. L. & Kim, J. Conserved sequence motifs for nucleoside triphosphate binding unique to turreted Reoviridae members and coltiviruses. *J. Virol.* **78,** 5528–5530 (2004).
15. Zhao, S., Liang, C., Hong, J. & Peng, H. Genomic sequence analyses of segments 1 to 6 of Dendrolimus punctatus cytoplasmic polyhedrosis virus. *Arch. Virol.* **148,** 1357–1368 (2003).
16. Sutton, G., Grimes, J. M., Stuart, D. I. & Roy, P. Bluetongue virus VP4 is an RNA-capping assembly line. *Nature Struct. Mol. Biol.* **14,** 449–451 (2007).
17. Stauber, N., Martinez-Costas, J., Sutton, G., Monastyrskaya, K. & Roy, P. Bluetongue virus VP6 protein binds ATP and exhibits an RNA-dependent ATPase function and a helicase activity that catalyze the unwinding of double-stranded RNA substrates. *J. Virol.* **71,** 7220–7226 (1997).
18. Kim, J., Parker, J. S., Murray, K. E. & Nibert, M. L. Nucleoside and RNA triphosphatase activities of orthoreovirus transcriptase cofactor μ2. *J. Biol. Chem.* **279,** 4394–4403 (2004).
19. Choi, K. H. & Rossmann, M. G. RNA-dependent RNA polymerases from Flaviviridae. *Curr. Opin. Struct. Biol.* **19,** 746–751 (2009).
20. Yang, C. *et al.* Cryo-EM structure of a transcribing cypovirus. *Proc. Natl Acad. Sci. USA* **109,** 6118–6123 (2012).
21. Yu, X., Jiang, J., Sun, J. & Zhou, Z. H. A. putative ATPase mediates RNA transcription and capping in a dsRNA virus. *Elife* **4,** e07901 (2015).
22. Luongo, C. L. *et al.* Loss of activities for mRNA synthesis accompanies loss of λ2 spikes from reovirus cores: an effect of λ2 on λ1 shell structure. *Virology* **296,** 24–38 (2002).
23. Patton, J. T., Jones, M. T., Kalbach, A. N., He, Y.-W. & Xiaobo, J. Rotavirus RNA polymerase requires the core shell protein to synthesize the double-stranded RNA genome. *J. Virol.* **71,** 9618–9626 (1997).
24. Mansell, E. A. & Patton, J. T. Rotavirus RNA replication: VP2, but not VP6, is necessary for viral replicase activity. *J. Virol.* **64,** 4988–4996 (1990).
25. Gridley, C. L. & Patton, J. T. Regulation of rotavirus polymerase activity by inner capsid proteins. *Curr. Opin. Virol.* **9,** 31–38 (2014).
26. McDonald, S. M. & Patton, J. T. Rotavirus VP2 core shell regions critical for viral polymerase activation. *J. Virol.* **85,** 3095–3105 (2011).
27. Starnes, M. C. & Joklik, W. K. Reovirus protein λ3 is a poly (C)-dependent poly (G) polymerase. *Virology* **193,** 356–366 (1993).
28. Reinisch, K. M., Nibert, M. L. & Harrison, S. C. Structure of the reovirus core at 3.6Å resolution. *Nature* **404,** 960–967 (2000).
29. Grimes, J. M. *et al.* The atomic structure of the bluetongue virus core. *Nature* **395,** 470–478 (1998).
30. Yu, X., Jin, L. & Zhou, Z. H. 3.88 Å structure of cytoplasmic polyhedrosis virus by cryo-electron microscopy. *Nature* **453,** 415–419 (2008).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported in part by grants from the National Institutes of Health (AI094386 and GM071940 to Z.H.Z.), NSFC (31172263 to J.S.) and NSFGD (S2013010016750 to J.S.). We acknowledge the use of instruments at the Electron Imaging Center for Nanomachines supported by UCLA and by instrumentation grants from NIH (1S10RR23057, 1S10OD018111) and NSF (DBI-1338135). We thank P. Ge for carrying out a Relion reconstruction without an initial model as an independent verification step, S. Schein and L. Wang for proof-reading the paper, and P. Afonine for model refinement.

Author Contributions Z.H.Z. supervised research; X.Z., X.Y. and Z.H.Z. designed and performed the experiments, analysed and interpreted data and wrote the paper; K.D. wrote programs, analysed data and prepared figures; W.C. built models; J.S. prepared reagents. All authors reviewed and finalized the paper.

Author Information 3D cryo-electron microscopy (cryoEM) density maps have been deposited in the Electron Microscopy Data Bank under the accession numbers EMD-6408 (3.3 Å averaged TEC in q-CPV), EMD-6404 (4.0 Å averaged TEC in t-CPV), EMD-6407 (3.9 Å full q-CPV), EMD-6405 (4.8 Å full t-CPV), EMD-6406 (5.1 Å asymmetric reconstruction of q-CPV by capsid-subtraction method) and EMD-6409 (filtered 22 Å q-CPV asymmetric reconstruction). The coordinates of atomic models of the TEC in q-CPV and t-CPV have been deposited in the Protein Data Bank under accession number 3JB6 and 3JB7, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to Z.H.Z. (Hong.Zhou@ucla.edu) or J.S. (cyfz@scau.edu.cn, for reagents).

METHODS

No statistical methods were used to predetermine sample size.

Sample preparation and cryoEM imaging. CPV particles were purified as described previously[30]. Purified polyhedra were treated at pH 10.8 with an alkaline solution (0.2 M Na_2CO_3-$NaHCO_3$) for 1 h, and then centrifuged at 10,000*g* for 40 min. The supernatant was collected and centrifuged at 80,000*g* for 60 min at 4 °C to pellet the CPV virions. The resulting pellet was directly re-suspended in the quiescent buffer (70 mM pH 8.0 Tris-Cl, 10 mM $MgCl_2$, 100 mM NaCl and 2 mM GTP). To prepare the transcribing CPV (t-CPV) particles, 30 µl purified CPV was incubated in a reaction buffer (70 mM Tris, pH 8.0, 10 mM $MgCl_2$, 100 mM NaCl, and 1 mM SAM + 2 mM GTP + 2 mM UTP + 2 mM CTP + 4 mM ATP) at 31 °C for 15 min, and then the reaction was stopped by quenching the reaction tubes on ice.

To prepare cryoEM grids, 2.5 µl of purified CPV sample was applied to a Quantifoil grid (2/2), blotted for 15 s with an FEI vitrobot in 100% humidity, and then plunged into liquid ethane. CryoEM images of the quiescent CPV (q-CPV) were collected in an FEI Titan Krios cryo electron microscope, operated at 300 kV with a nominal magnification of 49,000× (Extended Data Fig. 5g). The microscope was carefully aligned and electron beam tilt was minimized by a coma-free alignment procedure. Images were recorded on a Gatan K2 direct electron detection camera with the counting mode, and the pixel size was calibrated as 1.01 Å per pixel on the specimen using catalase crystals. The dose rate of the electron beam was set to ~8 e^- per pixel per s, and the image stacks were recorded at 4 frames s^{-1} for 3 s. The drift between frames in each image stack was corrected with the UCSF software[31], and the total 12 frames of each stack were merged to generate a final image with a total dose of ~25 e^- $Å^{-2}$. Contrast transfer function (CTF) parameters, including defocus values and astigmatism, were determined by CTFIND[32] (Extended Data Fig. 5g).

Sample grid preparation, cryoEM imaging and drift correction of frames for the transcribing CPV (t-CPV) were performed using the same procedure described above for q-CPV with the exception of the camera used. The t-CPV cryoEM images were recorded on a new Gatan K2 direct electron detection camera attached to a Gatan imaging filter (GIF Quanta) with a pixel size of 1.36 Å at the specimen scale (Extended Data Fig. 5g).

Asymmetric reconstruction based on original images. A total of 68,526 particles were selected for image processing using Frealign[33] and Relion[34]. The 2× binned data set was first processed using icosahedral symmetry with Frealign[33]. The centres of all particles were then fixed and used for the asymmetrical global search with Frealign using 4× binned data set starting at 20 Å resolution.

To generate an initial model, we placed the crystal structure of the MRV RdRP[2] under a previously obtained CPV capsid map[35] at the location corresponding to that in MRV capsid as previously reported[9] and imposed a tetrahedral symmetry (that is, with 4 three-fold axes, 3 two-fold axes and 12 asymmetric units), resulting in a montage map with an empty CPV capsid containing 12 RdRPs but without any VP4. This montage map was filtered to 30 Å resolution and used as the initial model for image processing with Frealign. After 9 iterations of global search and 2 iterations of refinement, the resolution of the density map was determined to be 3.9 Å. In the final map, only 3 RdRPs (numbers 8–10 in Fig. 1d) remained at the same locations as in the initial model with the tetrahedral symmetry.

The final map was reconstructed using the top 47,968 (70%) particles of the original unbinned data set. Averaging all TEC densities under different vertices was performed following the procedure described previously[36] to improve the density quality and the resolution. The effective resolution of the asymmetrical and averaged reconstructions were estimated to be 3.9 Å and 3.3 Å, respectively, based on the FSC ($\geq$0.143) and the correlation coefficient ($\geq$0.5) between the density map and atomic model calculated with Phenix (Extended Data Fig. 5g)[37,38]. These estimated resolutions are consistent with the observed structural features of the density maps (Fig. 2, Extended Data Fig. 5e and Supplementary Videos 3–8). The averaged map was filtered to the spatial frequency of 1/(3.3 Å) and sharpened with a reverse B-factor of –120 $Å^2$. This B-factor was chosen with a trial-and-error method based on the optimization of noise level, backbone density continuity, and emergence of side-chain densities.

Since there were no densities in the initial montage model at the VP4 locations, the emergence of VP4 densities in the map and the match of side-chain densities to those expected from the VP4 amino acid sequence (Fig. 2) provide strong internal controls for the validity of the high resolution cryoEM map. Consistent with this assessment, the locations of the RdRP in the final reconstruction are not only different from those in the initial montage model, but also are related by D3 symmetry instead of the tetrahedral symmetry in the initial model. Most convincingly, the density features in the final map agree with the CPV RdRP amino acid sequence but differ from that of the MRV RdRP used in the initial model.

In addition, we also performed independent reconstruction without using the model of the 12 MRV RdRPs, and obtained a nearly identical structure from the same data set. In this procedure, we first determined an icosahedral reconstruction without using any initial models. This icosahedral reconstruction was used to restrain refinement without symmetry (that is, symmetry operator is C1) to search for orientation around the 60 icosahedral-symmetry-related locations with Relion[34]. This independent result further validates our TEC structures.

To obtain the 3D structure of the transcribing particles, we low-pass filtered the above 3D map of q-CPV to 30 Å resolution and used it as the initial model. After 11 iterations of asymmetrical global search and 2 iterations of local refinement, the density map converged to a resolution of 4.8 Å, and the density quality of the TEC was further improved to ~4.0 Å resolution by aligning and averaging all TEC densities inside the asymmetric reconstruction (Extended Data Fig. 5d, f, g).

Asymmetric reconstruction using capsid-subtracted images. To improve the genome structure further, we used the following procedure to carry out asymmetric reconstruction of q-CPV with the same particle image data set but with capsid contribution subtracted. As illustrated in Extended Data Fig. 1, this procedure includes four stages: 1, capsid subtraction in raw particle (orange); 2, initial model generation (green); 3, asymmetric feature emergence in Relion[34] refinement (blue); 4, orientation selection (purple).

In the first stage (orange in Extended Data Fig. 1), we determined the orientation and centre parameters for each particle and obtained an icosahedral reconstruction with Frealign[33] from raw particles with an inverse B-factor of -40 $Å^2$ (Extended Data Fig. 1a, b). On the basis of these parameters, a CTF-corrected projection (Extended Data Fig. 1c) with empirical B-factor of 160 $Å^2$ was generated. Next, the capsid contribution to the images was removed by subtracting the 2D projection corresponding to the icosahedral orientation of each image as done before[39–42] with the following improvements. To subtract the contribution from the capsid accurately, we determined a scaling factor between capsid projection (Extended Data Fig. 1c) and each raw particle image (Extended Data Fig. 1a). The projection and raw images were both band-pass filtered between 1/400 $Å^{-1}$ and 1/29 $Å^{-1}$, then radially masked based on the inner and outer diameters of capsid to produce ring-shaped projections (Extended Data Fig. 1d) and raw (Extended Data Fig. 1e) images. The standard deviations of these ring-shaped images were calculated and used to normalize both the unmasked and masked (that is, ring-shaped) projections. The cross-correlation coefficient (0 to1) between the ring-shaped raw image and the normalized ring-shaped projection was computed and used as the probability factor measuring the contribution of capsid signal in the raw particle image. Each raw image was then subtracted by the unmasked projection multiplied by this probability factor to generate a capsid-subtracted particle (Extended Data Fig. 1f) for the following refinement. Particles with a probability factor less than 0.1 were not included in the subsequent analyses.

In the second stage (green in Extended Data Fig. 1), the map from the above Frealign asymmetric refinement (Extended Data Fig. 1g) was low-pass filtered to 60 Å resolution, masked with a 260 Å radius (Extended Data Fig. 1h), and used to refine the capsid-subtracted particle (Extended Data Fig. 1f) with Relion version 1.2. The Tau2_fudge value (T-factor) in Relion was set to 0.5. T-factor is an *ad hoc* value in Relion to tune refinement speed, and a value of 0.5 slowed down the refinement progression, thus ensuring the priority use of low resolution (up to 20 Å, such as dsRNA) data in the refinement. This refinement led to a reconstruction without the capsid (Extended Data Fig. 1i). This capsid-removed map has 12 TECs with D3 symmetry, which could be classified into two groups: the first group containing six better-resolved TECs close to the three-fold axis (polar), and the other group containing six less-resolved TECs near the equator (tropical), suggesting potential smear of density due to orientation mis-assignments or TEC flexibility/lower occupancy near the equator.

To eliminate potential orientation mis-assignments further, we next conducted the third stage of data processing (blue in Extended Data Fig. 1). We first low-pass filtered the capsid-removed reference (Extended Data Fig. 1i) to 32 Å resolution (Extended Data Fig. 1j) and used it to drive Relion refinement with the capsid-subtracted particles (Extended Data Fig. 1f). The T-factor used in this refinement is 0.1, only 2.5% of that used in Relion convention, thus ensuring slow progression of the refinement. Slower refinement provides time for asymmetrical feature to emerge. Relion global search was carried out with a 3.75° degrees angular interval, followed by local angular search with 1.875° interval and highly constrained translational search (0.7 pixel in range with 0.5 pixel interval). Asymmetrical RNA density feature with ten TECs emerged after ten iterations (Extended Data Fig. 1k). In our procedure, one way to prevent trapping into local minima in orientation assignment due to symmetric structural elements is to filter the current refinement result back to ~32 Å resolution and refine with T-factor of 0.1 again to remove residual symmetric feature from the working reference. This process is carried out iteratively.

To improve resolution of the 3D map further, we carried out the fourth stage for particle orientation selection (purple in Extended Data Fig. 1). From the

orientation of each particle determined in the high-resolution (~3 Å) icosahedral reconstruction (Extended Data Fig. 1b), we calculated 60 icosahedral-related orientation candidates. The task of the rest of the fourth stage of data processing is to select one out of these 60 orientation candidates to be the asymmetric orientation of the particle as done before[4,5,43]. To do this, we continued to run Relion refinement for 15 iterations using the above asymmetric map with 10 TECs (Extended Data Fig. 1k) as initial model and the orientation determined by each iteration was recorded, giving rise to 15 Relion orientations for each particle. For each of these 15 Relion orientations, we calculated its angular distances to the 60 icosahedral-related orientation candidates, and the icosahedral-related orientation candidate with the smallest angular distance was selected as the working orientation for that iteration, resulting in a total of 15 working orientations for each particle. The particle would be retained if 14 or all of its 15 working orientations are the same (that is, the selected orientation) and their averaged angular distance was less than 3 degrees. Otherwise, this particle will be discarded. This procedure yielded a total of 11,741 particles with selected orientation. The original raw images of these selected particles were combined to generate an asymmetric reconstruction using Frealign and the resolution was determined to be 5.1 Å.

As shown in Extended Data Fig. 2, this procedure was repeated by using a Gaussian ball to replace the capsid + TEC model (Extended Data Fig. 1g) in the initial model generation stage (green in Extended Data Fig. 1). The result is the same, confirming that our procedure was not influenced by the choice of initial model.

Atomic modelling and visualization. The atomic models of both RdRP and VP4 in the quiescent state were built with Coot[44] and refined with Phenix[38], as described previously[45].

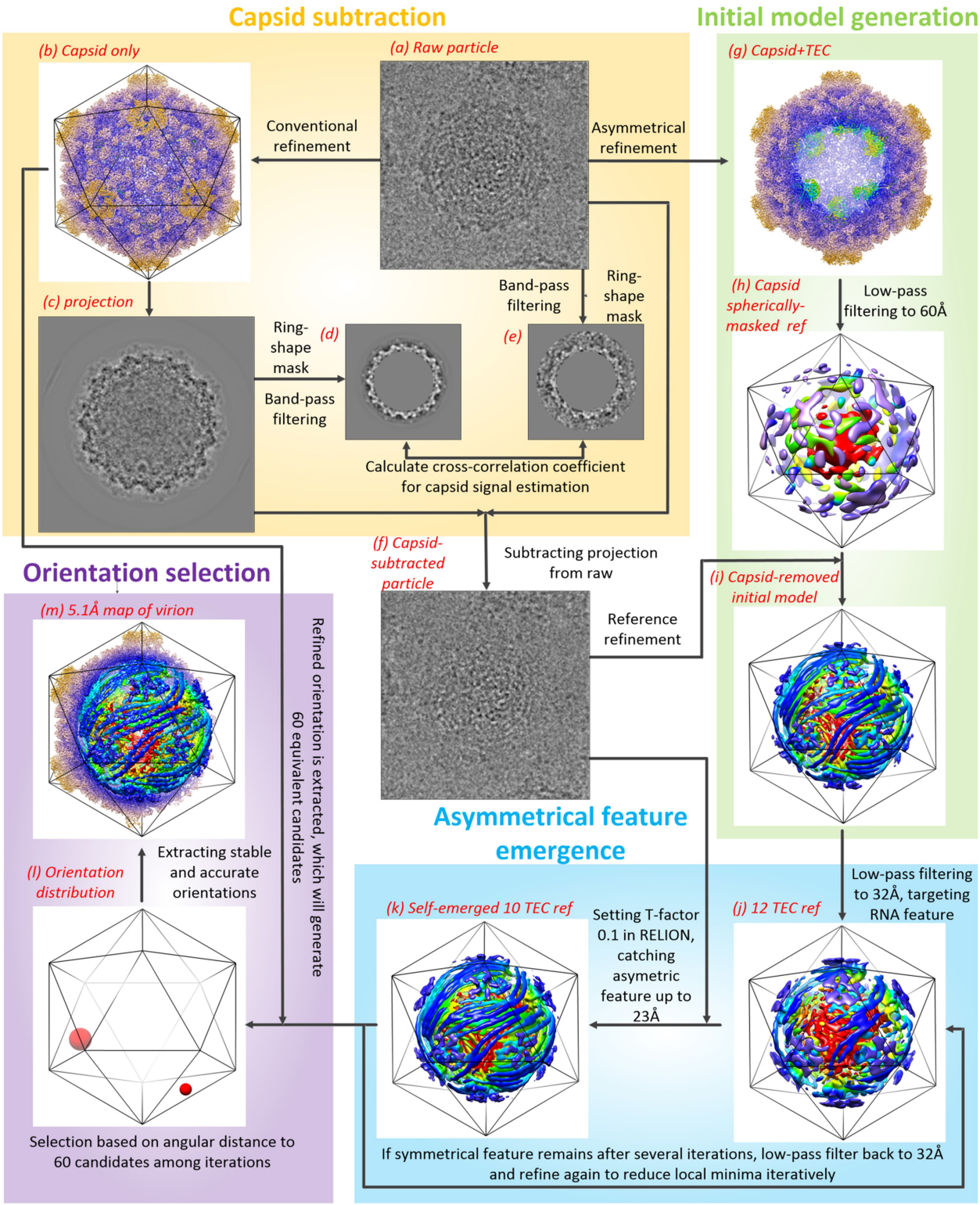
The atomic model of the VP4 structure was manually built with Coot. Because no homology models of VP4 previously existed, the Cα carbon backbone was constructed by matching the VP4 amino acid sequence to the density map. Once the correct placement of each residue was ensured, the backbone was converted to a purely alanine backbone by the function 'Mainchain', and mutated to the corresponding amino acids through the function 'Mutate Residue Range'. With the initial model now completed, the 'Density Fit Analysis' validation tool was used to screen for sequences of the model that did not fit the density. When identified, these sequences and the amino acids surrounding them were examined for any other possible conformations that would better fit the density. Owing to the high resolution of this structure, this was completed through the refinement tool 'Real Space Refine Zone', which optimizes the fit of the model to the mass density while preserving stereochemistry. Additionally, refinement was also performed based on the Ramachandran plot, an important indicator of three-dimensional protein structure that validates the torsion angles of a protein chain. In the Ramachandran plot, any residues with disallowed values were selected, and the stereochemistry of that residue along with its surrounding residues was optimized with the refinement tool 'Regularize Zone'. After ideal Ramachandran values were obtained (<1% outliers), the refinement function 'Rotamers' was used to select a rotamer that best fit the density.

The atomic model of the polymerase structure was also manually built with Coot. However, since an atomic model for the MRV polymerase was available in the Protein Data Bank (accession number 1MUK), this model was used as a template to assist with model building through the identification of the N terminus, C terminus and various secondary structures. Once the Cα carbon backbone was built by matching the polymerase amino acid sequence to the density map and mutated to the appropriate amino acids, the model was refined with 'Regularize Zone', 'Rotamers' and 'Real Space Refine Zone'. The model was validated with the Ramachandran plot and the function 'Density Fit Analysis'. The complex of VP4 and polymerase was then refined with Phenix, including the real space refinement[38].

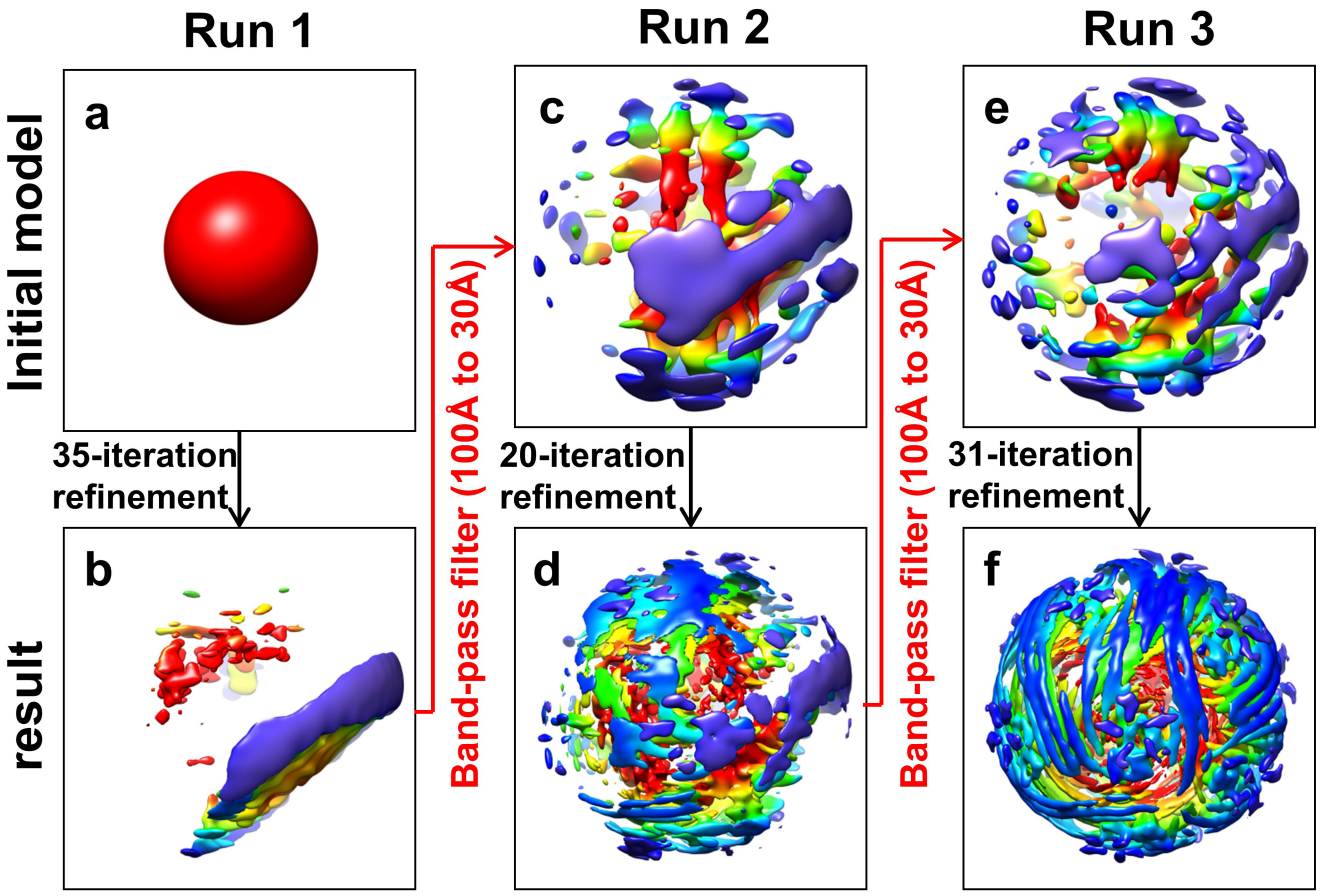
The atomic models of the transcribing state were built by fitting the atomic structures of RdRP and VP4 at quiescent state into the density, manually adjusting the changed residues with Coot[44], and refining the models with Phenix[38].

Visualization, segmentation of density maps, and generation of videos were done with UCSF Chimera[31].

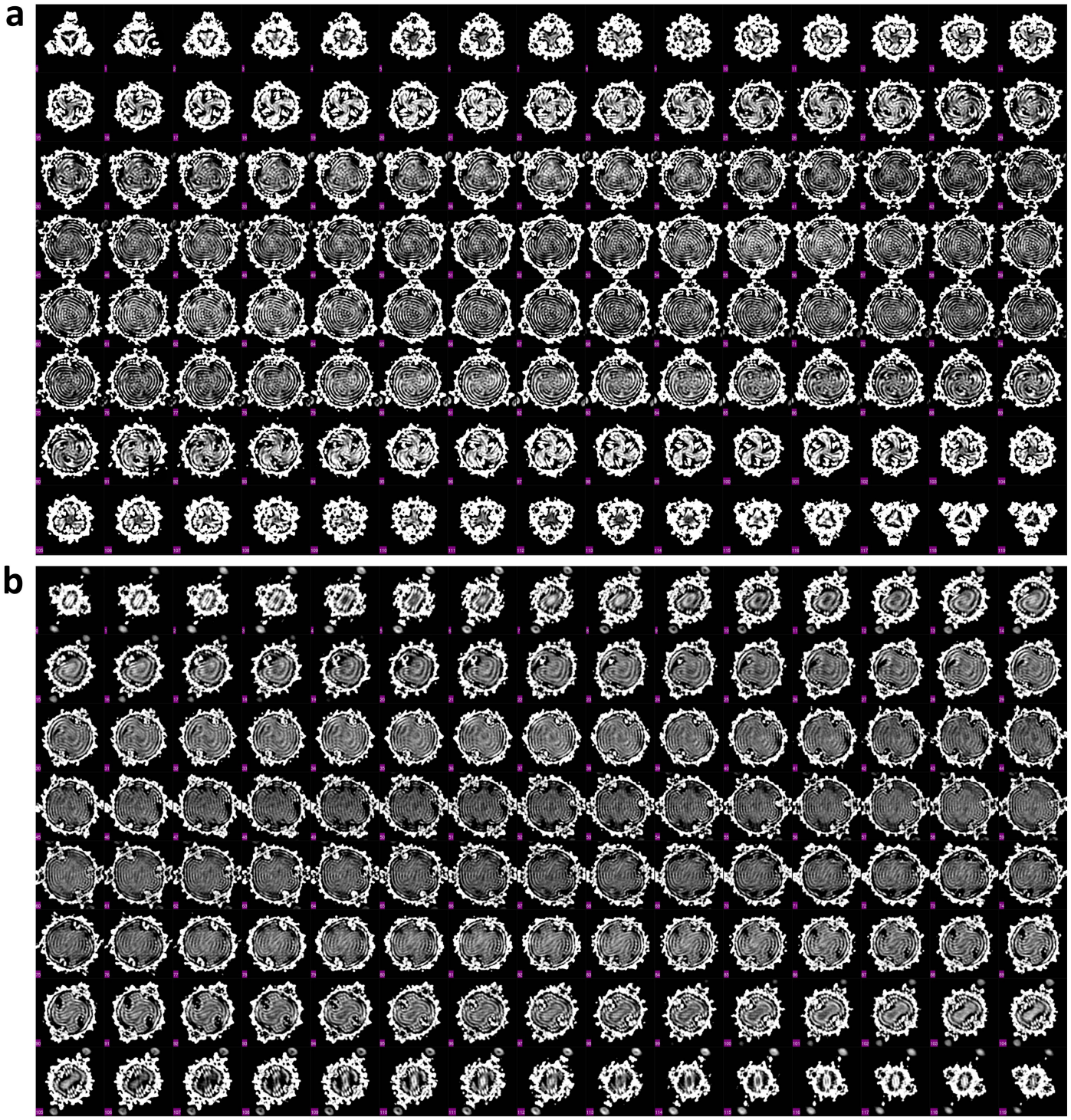
31. Pettersen, E. F. *et al.* UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25,** 1605–1612 (2004).
32. Mindell, J. A. & Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* **142,** 334–347 (2003).
33. Lyumkis, D., Brilot, A. F., Theobald, D. L. & Grigorieff, N. Likelihood-based classification of cryo-EM images using FREALIGN. *J. Struct. Biol.* **183,** 377–388 (2013).
34. Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180,** 519–530 (2012).
35. Yu, X., Ge, P., Jiang, J., Atanasov, I. & Zhou, Z. H. Atomic model of CPV reveals the mechanism used by this single-shelled virus to economically carry out functions conserved in multishelled reoviruses. *Structure* **19,** 652–661 (2011).
36. Zhang, X. *et al.* Near-atomic resolution using electron cryomicroscopy and single-particle reconstruction. *Proc. Natl Acad. Sci. USA* **105,** 1867–1872 (2008).
37. Wolf, M., Garcea, R. L., Grigorieff, N. & Harrison, S. C. Subunit interactions in bovine papillomavirus. *Proc. Natl Acad. Sci. USA* **107,** 6298–6303 (2010).
38. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66,** 213–221 (2010).
39. Huiskonen, J. T., Jaalinoja, H. T., Briggs, J. A., Fuller, S. D. & Butcher, S. J. Structure of a hexameric RNA packaging motor in a viral polymerase complex. *J. Struct. Biol.* **158,** 156–164 (2007).
40. Briggs, J. A. *et al.* Classification and three-dimensional reconstruction of unevenly distributed or symmetry mismatched features of icosahedral particles. *J. Struct. Biol.* **150,** 332–339 (2005).
41. Booy, F. *et al.* Liquid-crystalline, phage-like packing of encapsidated DNA in herpes simplex virus. *Cell* **64,** 1007–1015 (1991).
42. Zhang, Y., Kostyuchenko, V. A. & Rossmann, M. G. Structural analysis of viral nucleocapsids by subtraction of partial projections. *J. Struct. Biol.* **157,** 356–364 (2007).
43. Tao, Y. *et al.* Assembly of a tailed bacterial virus and its genome release studied in three dimensions. *Cell* **95,** 431–437 (1998).
44. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60,** 2126–2132 (2004).
45. Zhang, X. *et al.* A new topology of the HK97-like fold revealed in Bordetella bacteriophage by cryoEM at 3.5 Å resolution. *Elife* **2,** e01299 (2013).



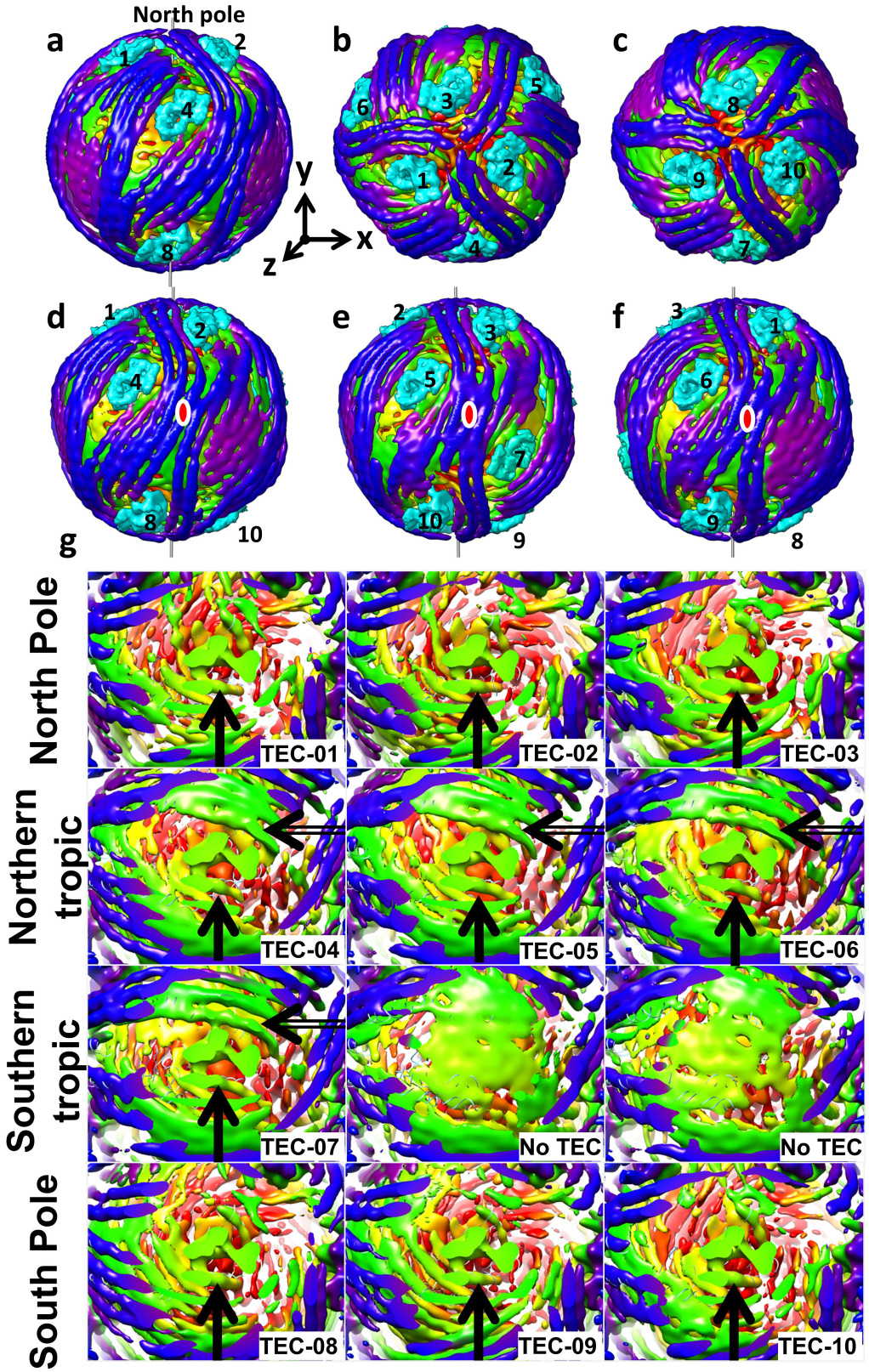
Extended Data Figure 1 | Illustration of the asymmetric reconstruction procedure using particles with the capsid density subtracted. See Methods for full explanation of panels **a**–**m**.



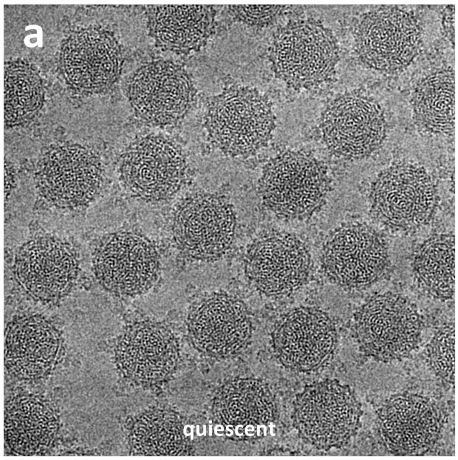
Extended Data Figure 2 | Validation of asymmetric reconstruction from capsid-subtracted images using a Gaussian ball as the initial model. Arrows linking **a** to **f** represent the progression of the procedure. The top panels (**a**, **c**, **e**) show the input model for each run and the bottom panels (**b**, **d**, **f**) show the output of each run.

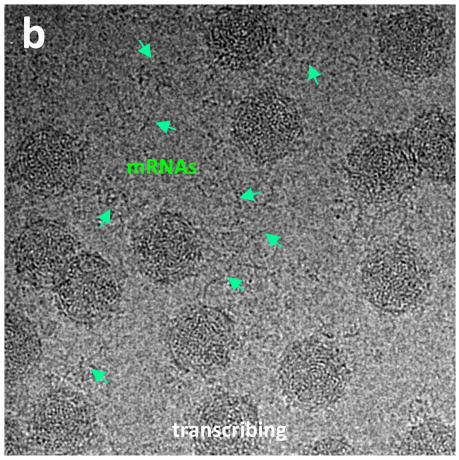


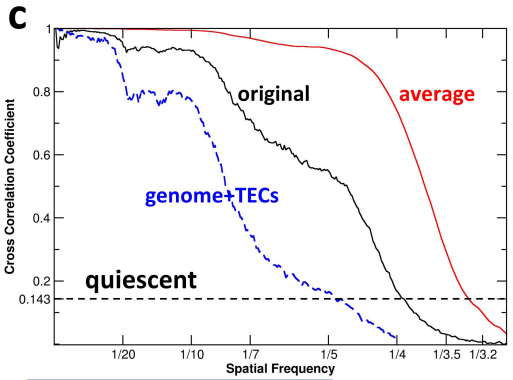
Extended Data Figure 3 | Sections of the q-CPV density map along the three-fold (that is, the earth axis) (a) **and two-fold** (b) **axes of the pseudo-D3 symmetry.** Note the lack of three-fold and two-fold symmetry in the RNA density in contrast to the perfect symmetry of the capsid shell proteins. Pixel size = 4.04 Å; clipped map size = 166 × 166 × 120 pixels.

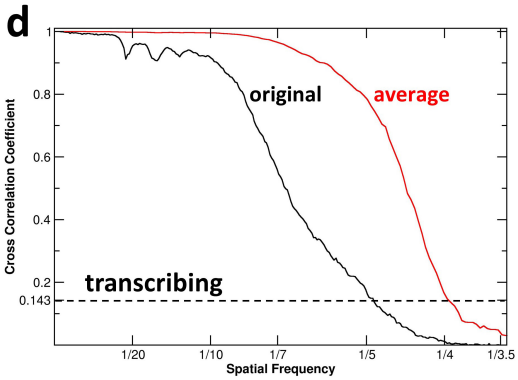


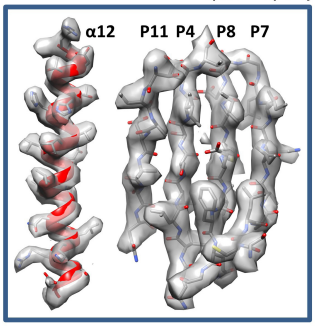
Extended Data Figure 4 | dsRNA density maps in the quiescent state. **a**, View of TEC + RNA densities with the same orientation of Fig. 1d. **b**, **c**, The same view as in **a** but rotated by +90° (**b**) or −90° (**c**) along *x* axis in panel **a** to view from either north (**b**) or south (**c**) poles. **d**–**f**, Three views from three two-fold axes on the equator, each is rotated by 120° along the *y* axis from each other. **g**, dsRNA density maps at the twelve vertices. TECs are arranged and numbered according to Fig. 1d. First row, TECs 1, 2, 3; second row, TECs 4, 5, 6; third row, TEC 7 and two unoccupied positions; fourth row, TECs 8, 9, 10. All TECs have a dsRNA segment bonded at the flange, each marked with a black arrow. Unlike the polar TECs, each tropical TEC (4–7) is surrounded by an extra density rod (open arrow).

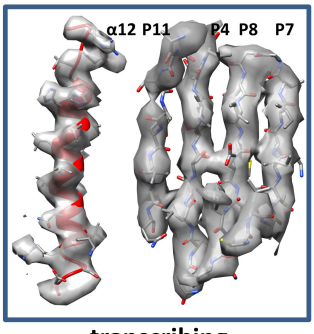








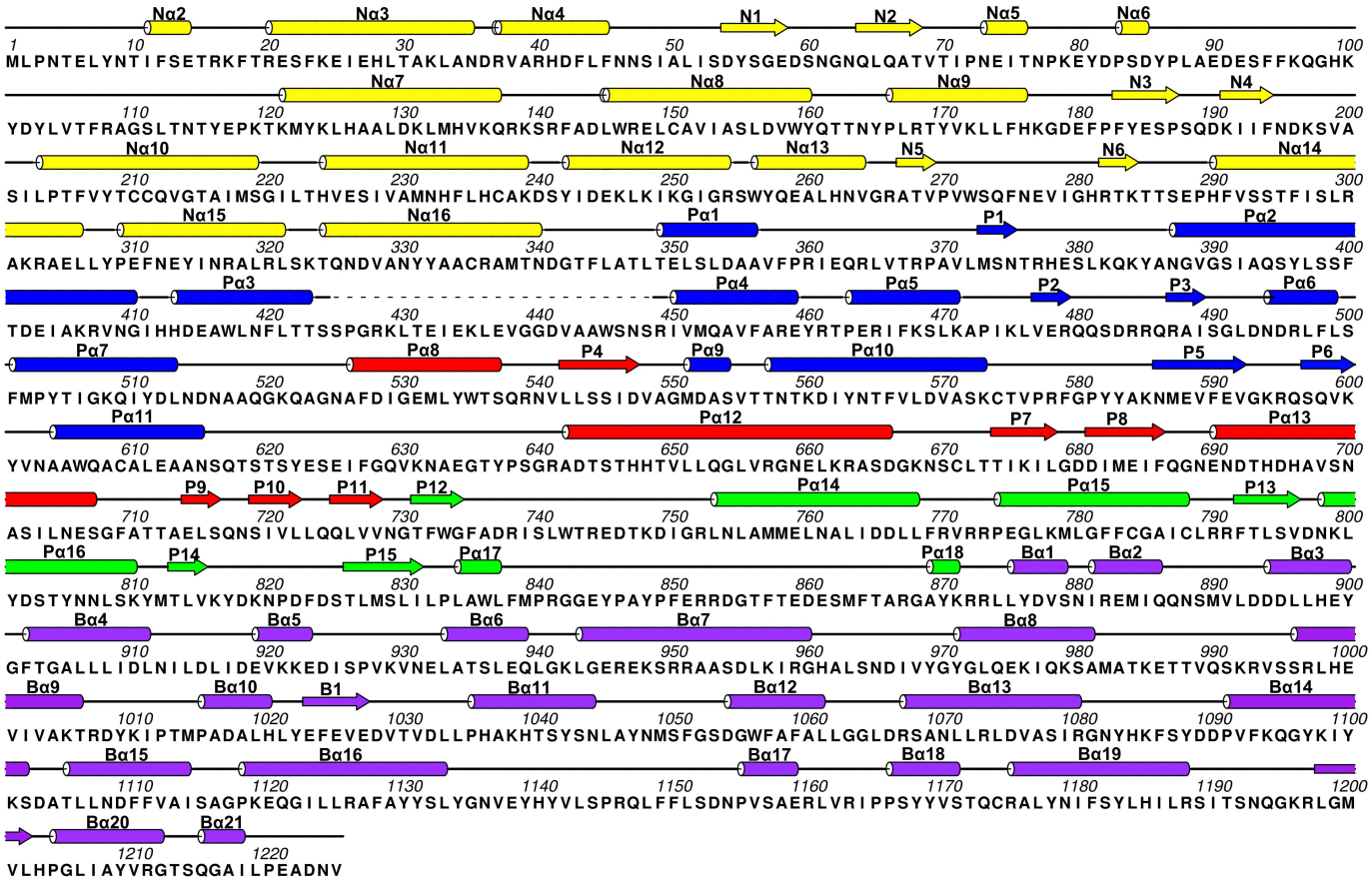




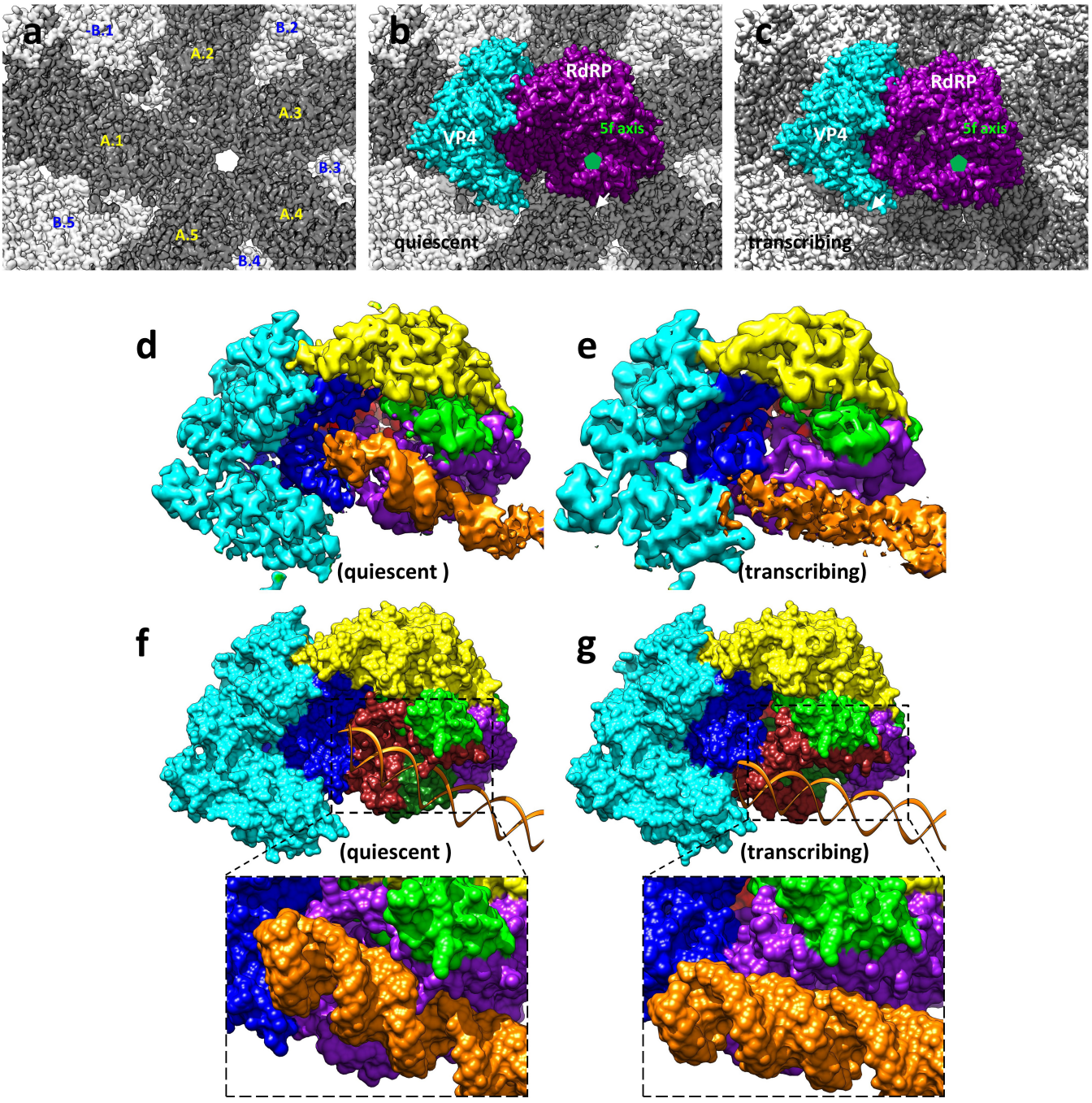
g

	Quiescent state	Transcribing state
Sample buffer (pH8.0)	70mM pH 8.0 Tris-Cl, 10 mM $MgCl_2$, 100 mM NaCl and 2 mM GTP	70mM Tris, pH 8.0, 10 mM $MgCl_2$, 100 mM NaCl, and 1 mM SAM+ 2mM GTP+2mM UTP+ 2mM CTP+4 mM ATP
Microscope	FEI Titan Krios	FEI Titan Krios
Voltage (kV)	300	300
Camera	Gatan K2 (counting mode)	Gatan K2 (counting mode)
Energy-filter	None	GIF Quantum Energy Filter
Nominal Magnification	49,000x	105,000x
Å/pixel	1.01	1.36
Underdefocus range (μm)	0.6~4.51	0.89~3.4
Total Dose (electrons/Å^2)	~25	~40
Frames per micrograph	12	40
Micrographs exposed	4385	4907
Particles selected	68526	81887
Resolution (Å) (FSC≥0.143)	3.3 (averaged)	4.0 (averaged)
Model refinement with phenix		
R_{work} (overall)	0.18 (50-3.3Å)	0.25 (50-4.0Å)
R_{free} (overall)	0.17 (50-3.3Å)	0.23 (50-4.0Å)
R_{work} (best resolution zone)	0.37 (3.6-3.3Å)	0.35 (4.6-4.0Å)
R_{free} (best resolution zone)	0.36 (3.6-3.3Å)	0.32 (4.6-4.0Å)
Ramachandran plot values		
Most favored	97.57%	95.91%
Generously allowed	2.38%	3.91%
Disallowed regions	0.06%	0.18%

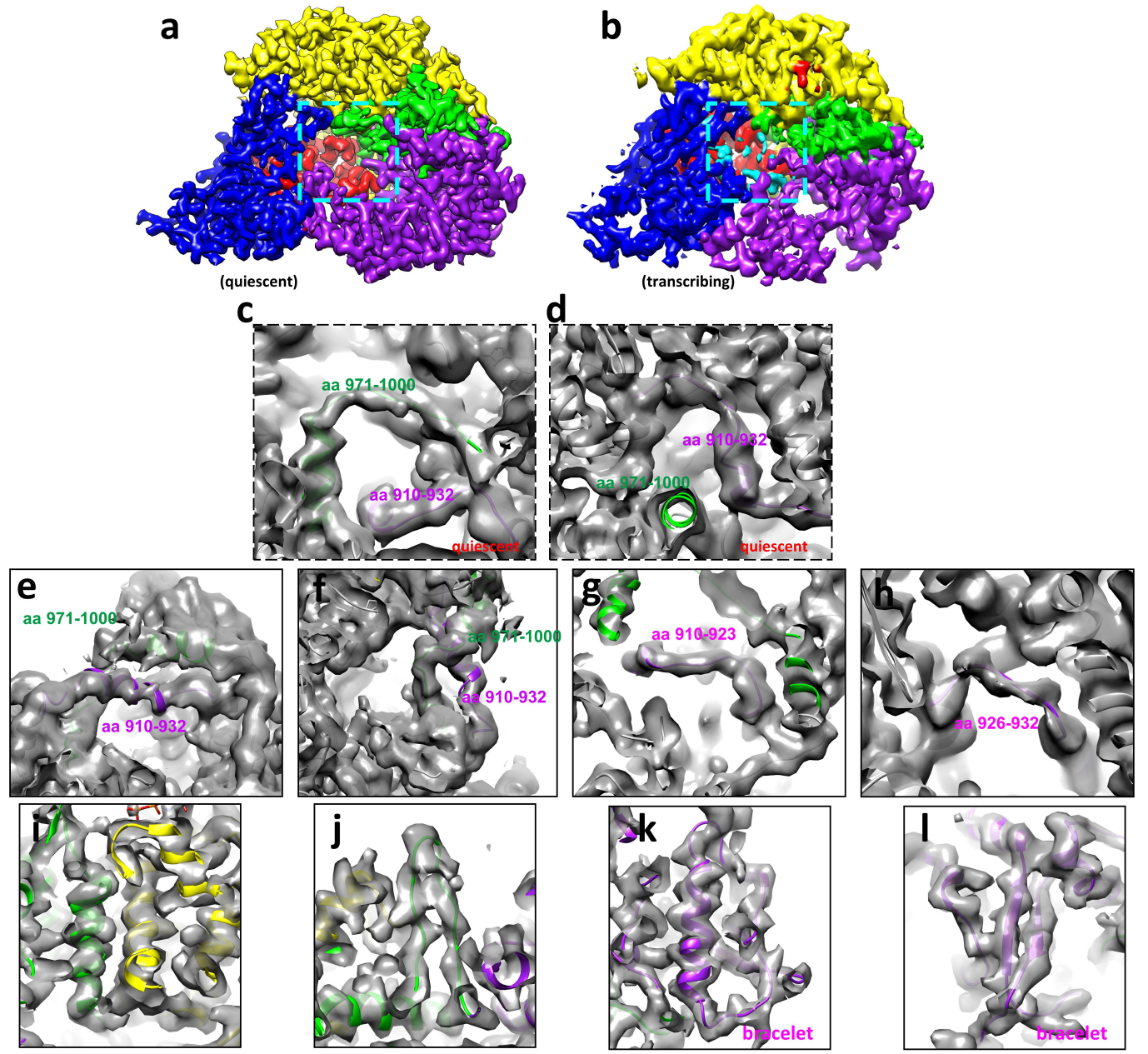
Extended Data Figure 5 | CryoEM reconstructions of CPV in the quiescent and transcribing states. a, **b**, CryoEM images of CPV particles in quiescent (**a**) and transcribing (**b**) states. These images were obtained by aligning and averaging frames in direct electron counting image stacks. Fibre-like nascent mRNAs are visible over background in **b** (marked by green arrows), while the background in **a** is clean. **c**, **d**, Fourier shell correlation coefficients (FSCs) as a function of spatial frequency between two half maps for reconstructions in the quiescent (**c**) and transcribing (**d**) states. The black and red lines represent FSCs for the asymmetrical reconstructions of capsid + genome and the locally averaged TEC densities, respectively. The effective resolutions of the local averaged maps are ~3.3 Å (**c**) and ~4.0 Å (**d**) resolution (FSC≥0.143) for maps in the quiescent and transcribing, respectively. **e**, **f**, CryoEM densities (grey surface representations) superimposed with atomic models (ribbons and sticks) for the quiescent (**e**) and transcribing (**f**) states. The α-helix (Pα12) and the four-stranded β-sheet (P4, P7–8 and P11) in **e** and **f** are both from the palm subdomain of the polymerase domain at 3.3 Å (**e**) and ~4.0 Å (**f**) resolutions. **g**, Statistics of CPV reconstructions and atomic model refinement.



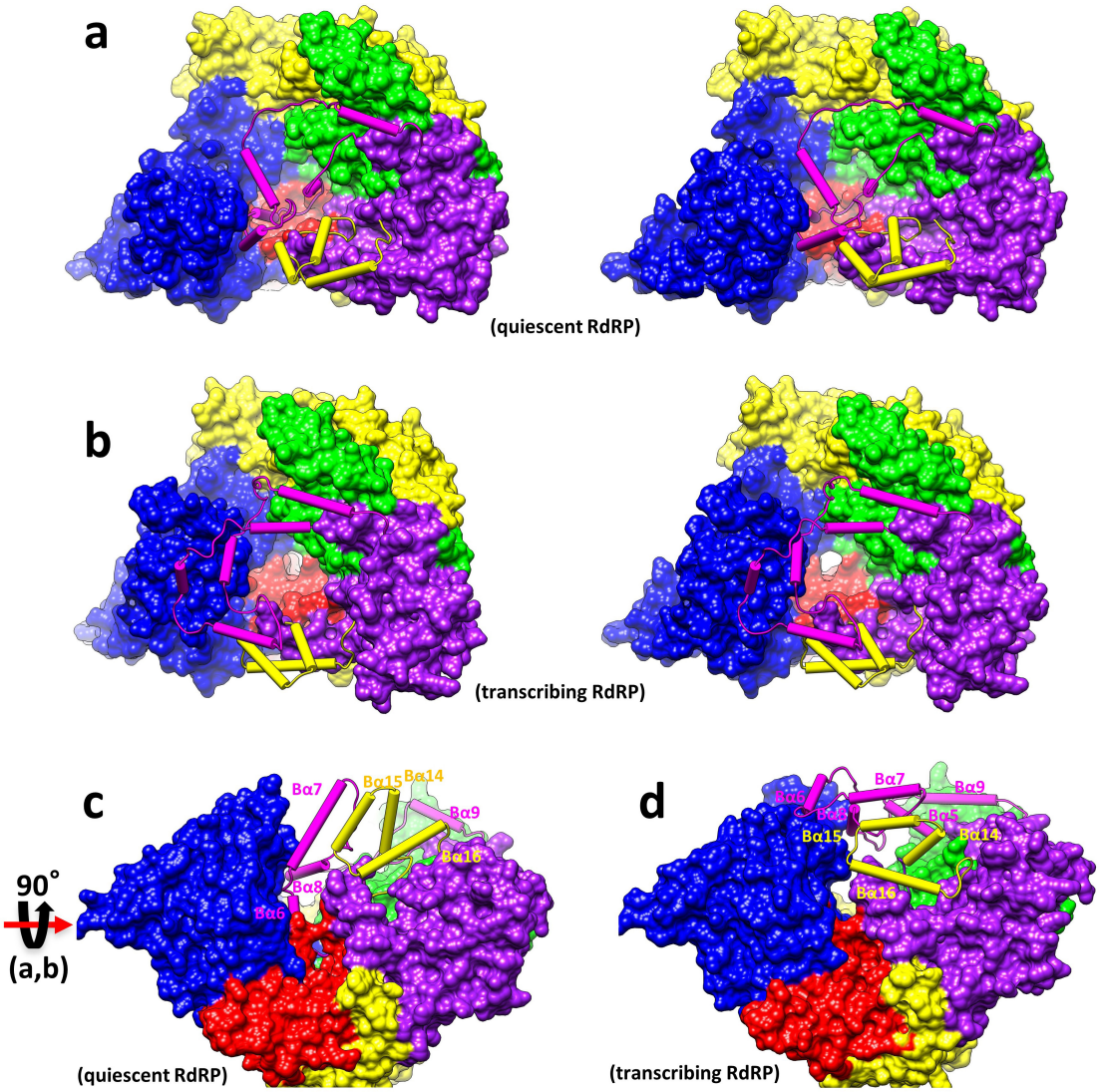
Extended Data Figure 6 | Sequence and secondary structure assignment of CPV RdRP in the quiescent state. α-helices were marked by cylinders, β-strands by arrows, loops by thin lines, and the flexible tip domain by dashed lines. The colour scheme is the same as Fig. 3a.



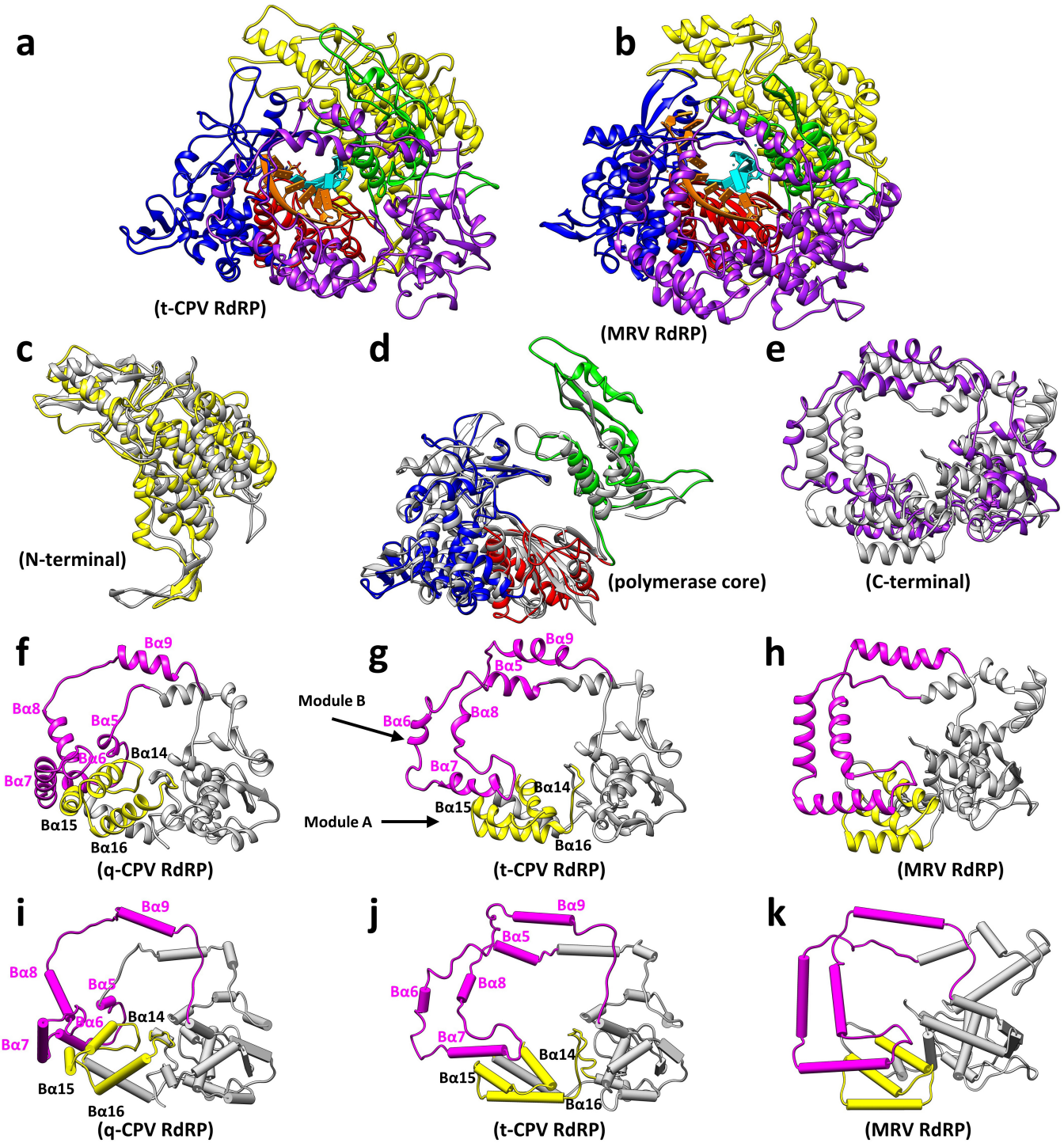
Extended Data Figure 7 | The RdRP-bound dsRNA in the quiescent and transcribing states. a–**c**, Location of a TEC on the inner surface of the capsid shell in the quiescent and transcribing states. The inner surface of the CPV capsid (**a**) with 10 CSPs labelled (CSP A.1/B.1 to CSP A.5/B.5). **b**, **c**, Position of a TEC on the inner surface of capsid in the quiescent (**b**) and transcribing (**c**) states. VP4 and RdRP are coloured cyan and purple, respectively. An icosahedral five-fold axis is indicated with a small green pentagon. **d**, **e**, CryoEM densities of TEC and dsRNA (orange) in the quiescent (**d**) and transcribing (**e**) states. **f**, **g**, Models of TEC (surface representation) and dsRNA (ribbons) in the quiescent (**f**) and transcribing (**g**) states. Close-up views show the bound dsRNA (surface representation) on RdRP in the quiescent state (**f**) and its detachment in the transcribing state (**g**). VP4 is coloured cyan and the RdRP is coloured as in Fig. 3a. All surfaces displayed in this figure were rendered from models, except for the density maps of RdRP + dsRNA in **d**, **e**.



Extended Data Figure 8 | Tracing amino acid residues 910–932 and 971–1000 of module B of the bracelet domain of RdRP in the quiescent and transcribing states. a, **b**, CryoEM densities of RdRP in the quiescent (**a**) and transcribing (**b**) states. The locations of the residues 910–932 and 971–1000 are indicated with cyan boxes in **a** and **b**. Owing to their flexibility, these residues are not readily visible when displayed as in **a** and **b** but become visible when the maps are filtered to a lower resolution (for example, 4.5 Å resolution) as in **c**–**f**. The colour scheme of domains/subdomains is the same as in Fig. 3a. **c**, Trace of the residues 971–1000 (green) and 910–932 (purple) of module B of the bracelet domain of RdRP in the quiescent state. **d**, The same as **c** but in a different view. **e**, Trace of the residues 971–1000 (green) and 910–932 (purple) of module B of the bracelet domain of RdRP in the transcribing state. **f**, The same as **e** but in a different view to show the unambiguous trace of the two peptide fragments. **g**, **h**, Trace of the residues 910–923 (**g**) (purple) and 926–932 (**h**) (purple) of the bracelet domain of RdRP in the transcribing state, showing the unambiguous trace of the two peptide fragments. **i**, **j**, CryoEM densities (grey) and model (ribbon) of RdRP in the transcribing state, showing α-helices (**i**) and a β-hairpin (**j**). The colour scheme of domains/subdomains is the same as in Fig. 3a, **k**, **l**, Trace of the residues of the bracelet domain of RdRP in the transcribing state, showing a α-helix (**k**) and a β-sheet (**l**).



Extended Data Figure 9 | Stereo and rotated views of Fig. 4a, b. a, **b**, Stereo views of modules A (yellow cylinders and loops) and B (purple cylinders and loops) of the bracelet domain of RdRP in the quiescent **(a)** and transcribing **(b)** states. **c**, **d**, Same as in **a**, **b**, but rotated around the *x* axis by 90°. All surfaces displayed in this figure were rendered from models.



Extended Data Figure 10 | Comparisons of RdRPs from CPV and MRV. **a**, **b**, CryoEM *in situ* structure of the RdRP in t-CPV (**a**) and crystal structure of the MRV RdRP (**b**), both containing a RNA duplex in the active site. **c**–**e**, Superposition of domains of RdRPs from t-CPV (colour) and MRV (grey): N-terminal (**c**), polymerase (**d**) and bracelet (**e**) domains. **f**–**h**, Comparisons of modules A (yellow) and B (magenta) of the bracelet domain of RdRPs from q-CPV (**f**), t-CPV (**g**) and MRV (**h**). **i**–**k**, The same as in **f**–**h**, but with helices shown as cylinders, as in Fig. 4a, b.

LETTER

doi:10.1038/nature15760

Foreign DNA capture during CRISPR–Cas adaptive immunity

James K. Nuñez[1*], Lucas B. Harrington[1*], Philip J. Kranzusch[1,2], Alan N. Engelman[3,4] & Jennifer A. Doudna[1,2,5,6,7,8]

Bacteria and archaea generate adaptive immunity against phages and plasmids by integrating foreign DNA of specific 30–40-base-pair lengths into clustered regularly interspaced short palindromic repeat (CRISPR) loci as spacer segments[1–6]. The universally conserved Cas1–Cas2 integrase complex catalyses spacer acquisition using a direct nucleophilic integration mechanism similar to retroviral integrases and transposases[7–13]. How the Cas1–Cas2 complex selects foreign DNA substrates for integration remains unknown. Here we present X-ray crystal structures of the *Escherichia coli* Cas1–Cas2 complex bound to cognate 33-nucleotide protospacer DNA substrates. The protein complex creates a curved binding surface spanning the length of the DNA and splays the ends of the protospacer to allow each terminal nucleophilic 3′-OH to enter a channel leading into the Cas1 active sites. Phosphodiester backbone interactions between the protospacer and the proteins explain the sequence-nonspecific substrate selection observed *in vivo*[2–4]. Our results uncover the structural basis for foreign DNA capture and the mechanism by which Cas1–Cas2 functions as a molecular ruler to dictate the sequence architecture of CRISPR loci.

CRISPR loci are defined by repetitive elements that are separated by similarly sized spacer sequences acquired from foreign DNA during the adaptation stage of CRISPR–Cas adaptive immunity[6,14]. CRISPR transcripts generated from the loci assemble with Cas proteins to detect and cleave foreign nucleic acids bearing sequence complementarity to the spacer segment[1,5,15–19]. In *E. coli*, expression of the Cas1–Cas2 protein complex triggers acquisition of new 33-base-pair (bp) spacers at the A/T-rich leader end of the CRISPR locus[7–10,20]. How the Cas1–Cas2 complex selects 33-bp protospacers of variable sequences and activates the 3′-OH ends for integration remains unknown. As the Cas1–Cas2 complex is sufficient to initiate spacer acquisition and adaptation of the CRISPR–Cas immune system, we hypothesized that the protein complex alone must provide the structural basis for the unknown mechanism of spacer length determination.

To determine how protospacer variation influences the efficiency of Cas1–Cas2-mediated spacer acquisition, we used an *in vitro* integration assay to test versions of a 33 bp sequence with constant overall length but different 3′ single-stranded overhang lengths[12]. The protospacer sequence is derived from the M13 bacteriophage genome and is highly acquired into the *E. coli* CRISPR locus after infection[8]. Unexpectedly, protospacers with overhanging 3′ nucleotides are strongly preferred by the Cas1–Cas2 complex over a completely double-stranded 33 bp protospacer (Fig. 1a and Extended Data Fig. 1a, b). Single-stranded DNA and substrates with 5′ overhangs are poor substrates for integration, highlighting the ability of Cas1–Cas2 to select specific DNA substrates before integration[12]. The most preferred protospacer DNA for *in vitro* integration consists of five overhanging nucleotides on each 3′ end (Extended Data Fig. 1). To determine the molecular basis of Cas1–Cas2 protospacer capture, we assembled Cas1–Cas2 complexes with the preferred protospacer substrate and determined crystal structures of the complex in the presence and absence of Mg^{2+} at 3.0 Å and 3.2 Å resolutions, respectively (Extended Data Fig. 2 and Extended Data Table 1).

The structures reveal a hexameric protein architecture comprising four copies of Cas1 and two copies of Cas2, in which the protospacer spans the central Cas2 dimer and terminates within individual Cas1 subunits on each end of the complex (Fig. 1b). Structural superposition of the Cas1–Cas2 complex with and without bound DNA reveals a DNA-induced change in Cas1 subunit orientation in which each Cas1 dimer rotates ~10° in opposing directions against the central Cas2 hub (Extended Data Fig. 3a, b). Cas1–Cas2 protospacer capture positions each single-stranded protospacer 3′ end within a channel leading directly to a Cas1 active site. Simulated annealing omit maps show clear electron density for the double-helical region and the five-nucleotide overhangs on each end of the protospacer (Extended Data Fig. 4a–c). The constrained protein channel guiding each DNA strand from its

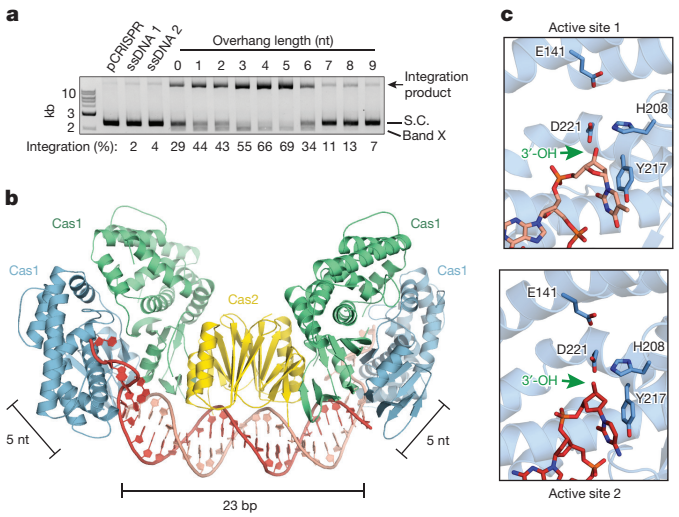


Figure 1 | Overall architecture and active site positioning of 3′-OH nucleophile. a, A representative agarose gel of *in vitro* integration reactions using increasing lengths of 3′ single-strand (ss) protospacer DNA overhangs. Per cent integration values are the average of three independent experiments. kb, kilobases; nt, nucleotide; S.C., supercoiled pCRISPR; Band X, relaxed pCRISPR byproduct (ref. 12). **b**, The overall architecture of Cas1–Cas2 bound to protospacer DNA. The line segments indicate the length of the DNA, spanning a total of 33 nucleotides. **c**, Stick configurations of the two Cas1 active sites (blue subunits in **b**) that coordinate the nucleophilic 3′-OH ends of the protospacer (green arrow). Supplementary Information contains the full image for **a**.

[1]Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, California 94720, USA. [2]Howard Hughes Medical Institute, University of California, Berkeley, Berkeley, California 94720, USA. [3]Department of Cancer Immunology and Virology, Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA. [4]Department of Medicine, Harvard Medical School, Boston, Massachusetts 02115, USA. [5]Department of Chemistry, University of California, Berkeley, Berkeley, California 94720, USA. [6]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA. [7]Innovative Genomics Initiative, University of California, Berkeley, Berkeley, California 94720, USA. [8]Center for RNA Systems Biology, University of California, Berkeley, Berkeley, California 94720, USA.
*These authors contributed equally to this work.

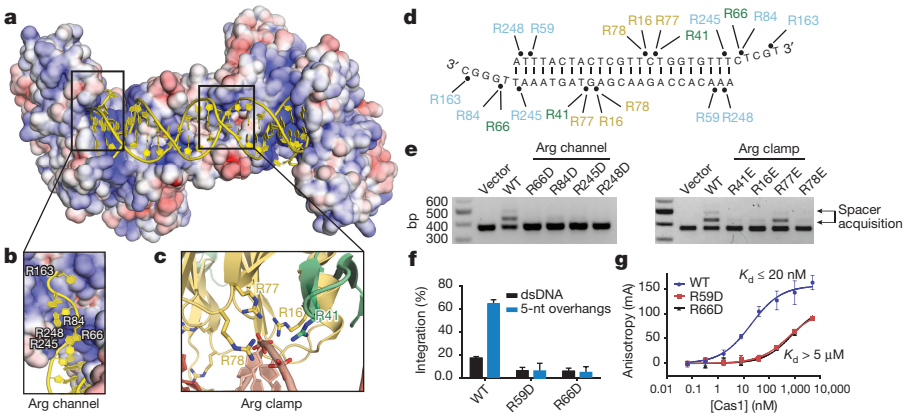


Figure 2 | Coordination of protospacer DNA within the complex. **a**, Electrostatic potential surface representation of the Cas1–Cas2 complex with the protospacer shown in yellow. **b**, Close up of the arginine channel that stabilizes the ssDNA overhang. **c**, Stick configuration representation of arginine clamp residues that coordinate the protospacer duplex region. **d**, Map of amino acid residues that coordinate the protospacer phosphodiester backbone (black dots). Residue colours indicate Cas1–Cas2 protomers from Fig. 1b. **e**, Agarose gels of *in vivo* spacer acquisition assays of arginine channel and clamp mutant proteins. WT, wild type. **f**, Plot of per cent *in vitro* integration of either double-stranded DNA (dsDNA; black) or 5-nucleotide (nt) overhang (blue) protospacers with wild-type Cas1, Cas1(R59D) or Cas1(R66D) complexed with Cas2. **g**, Fluorescence polarization binding assays of a 5-nucleotide overhang protospacer with the same mutants in **f** complexed with Cas2. The calculated relative binding affinities (K_d) are indicated. Error bars represent the standard deviation of three independent experiments. Data in panel **e**–**g** are results of at least three biological replicates. Supplementary Information contains the full images for **e**.

double-helical region to the single-strand-accommodating Cas1 active site explains the specificity of Cas1–Cas2 for five-nucleotide 3′ overhang substrates (Fig. 1a and Extended Data Fig. 1). Two of the four Cas1 subunits, coloured green in Fig. 1b, are not occupied with the protospacer 3′ ends and are probably non-catalytic, since the 3′-OH nucleophile and the scissile phosphodiester bond of the target DNA must be in the same active site for direct nucleophilic integration.

In the active sites, the 3′ terminal base is involved in a stacking interaction with Y217 that positions the nucleophilic 3′-OH ends of the protospacer near the conserved metal-binding residues E141, H208 and D221 (Fig. 1c). Although we cannot assign density for Mg^{2+} in the active sites, these three residues have been shown previously to coordinate a Mn^{2+} ion in the active site of Cas1 from *Pseudomonas aeruginosa*[21]. Furthermore, alanine mutations at these positions disrupt *in vivo* spacer acquisition[7,8,10]. Thus, the observed positioning of the 3′-OH nucleophiles and catalytic residues probably represents the active configuration of the nucleoprotein complex immediately before spacer integration.

All interactions between Cas1–Cas2 and protospacer DNA involve coordination of the phosphate backbone rather than base-specific contacts, consistent with the variable sequence selection of protospacers that is essential for resistance to diverse foreign sequences[2–4]. Two central regions of the Cas1–Cas2 complex, which we term the 'arginine clamp' and the 'arginine channel', stabilize the protospacer (Fig. 2a–d). The arginine clamp interacts with the middle of the duplex region where four Arg residues coordinate each DNA strand: Cas1 R41 and Cas2 R16, R77 and R78 (Fig. 2c). Reverse charge mutations of Cas1 R41 and Cas2 R16 and R78 drastically reduce spacer acquisition *in vivo*, whereas the Cas2 R77E mutant functions similar to wild-type Cas2 (Fig. 2e). Thus, Cas1 R41, Cas2 R16 and R78 are the key constituents of the arginine clamp. The contribution of Cas2 to protospacer DNA binding supports the previous hypothesis that the main function of Cas2 is to form a non-catalytic scaffold within the Cas1–Cas2 complex[10].

Cas1 residues R66, R84, R245 and R248 line the arginine channel that stabilizes the junction where the duplex region terminates and the single-stranded DNA overhang enters the active site. Reverse charge mutations of each arginine lining the arginine channel disrupts spacer acquisition *in vivo* (Fig. 2e). In addition, purified Cas1 R59D or R66D proteins complexed with wild-type Cas2 are highly defective in integrating 33-bp duplex or five-nucleotide overhang protospacer substrates *in vitro* (Fig. 2f). Fluorescence polarization assays demonstrate that the mutant complexes exhibit dramatically reduced affinity for protospacer DNA, highlighting the critical role of this part of the Cas1–Cas2 complex for protospacer capture and complex stability (Fig. 2g).

The Cas1–Cas2–DNA crystal structures uncover a protein wedge that terminates the protospacer double-stranded DNA region and allows single-stranded DNA overhangs to enter the arginine channel. A stacking interaction of the 5′ terminal base (adenine 6 in Fig. 3a, b) with Y22 of Cas1 stabilizes protospacer duplex unwinding, directing each single-stranded 3′ overhang to sharply bend ~90° away from the duplex and into the active site channel (Fig. 3b). A mutation of Y22 to alanine reduces spacer acquisition *in vivo*, whereas a phenylalanine mutation has near wild-type levels of acquisition, consistent with a specific role for Cas1 Y22 base-stacking in protospacer strand splaying (Fig. 3c).

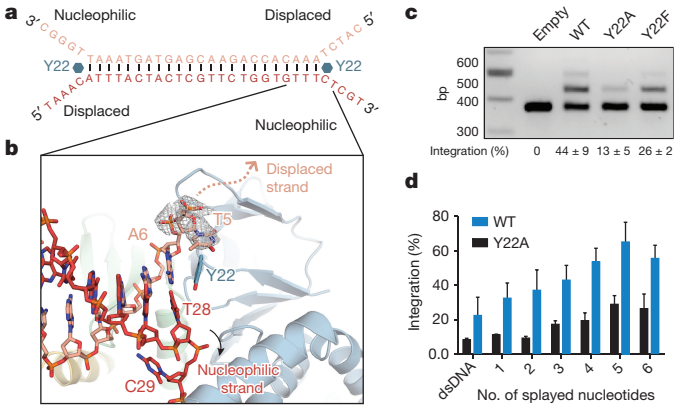


Figure 3 | Mechanism of protospacer DNA end separation. **a**, The 5-nucleotide splayed protospacer sequence used for crystallization to determine the trajectory of the displaced non-nucleophilic strand. Cas1 Y22, involved in base stacking at the fork, is shown in blue. **b**, Close up of the DNA fork showing the base stacking interaction of Y22 with the terminal adenine nucleotide of the non-nucleophilic strand. The nucleotides are numbered from 5′ to 3′ of each DNA strand shown in **a**. The grey mesh shows the $2F_o - F_c$ density contoured at 2.2σ of the first ejected nucleotide of the displaced strand. The arrows indicate the opposite trajectories of each strand. **c**, Agarose gel of *in vivo* acquisition assay of co-expressed wild-type (WT) Cas1 or the indicated Cas1 mutant with Cas2. Quantification is the mean of three independent experiments ± standard deviation. **d**, Plot of per cent integration of increasing number of splayed nucleotides at the protospacer ends using wild-type Cas1 (blue) or Cas1(Y22A) (blue) complexed with Cas2. Error bars represent the standard deviation of three independent experiments. Supplementary Information contains the full image for **c**.

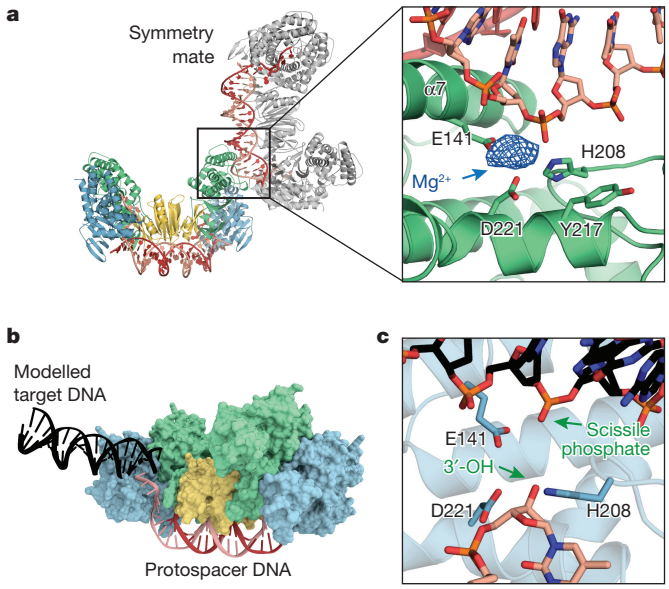


Figure 4 | Model of protospacer DNA integration. a, View of crystal packing from a symmetry mate complex (grey) showing coordination of the symmetry DNA along a Cas1 active site. The inset is a magnified view of the coordination of the phosphodiester backbone with metal-binding residues E141, H208 and D221. The mesh represents a $F_o - F_c$ density for a Mg^{2+} ion, contoured at 2.2σ. **b**, **c**, Model of protospacer DNA integration into target DNA (black) and positioning of the scissile phosphate (green arrow) and the 3′-OH nucleophile in the Cas1 active site.

Sequence alignment of representative Cas1 proteins in type I CRISPR systems reveals that Y22 is not universally conserved in other bacteria, suggesting that additional or different Cas1 residues may stabilize the splayed ends in other CRISPR–Cas systems (Extended Data Fig. 5).

The observed stacking interaction raises the possibility that fully duplexed protospacers are separated by Cas1 Y22, thereby displacing the 5′ end of the duplex, which we term the non-nucleophilic strand, from the nucleophilic strand carrying the 3′-OH. DNA transposases and retroviral integrases also utilize end fraying to isolate the reactive DNA strands for chemistry within enzyme active sites[22–24]. To test this potential activity of Cas1–Cas2, we introduced an increasing number of mismatches at the ends of the 33 bp protospacer to disrupt end base pairing and assayed their potential for *in vitro* integration (Fig. 3d and Extended Data Fig. 6a, b). Similar to the 3′ overhang substrates, the 4- and 5-nucleotide frayed ends are highly preferred, presumably due to the lower energy required for capture of these substrates compared to perfectly duplexed ends (Fig. 3d). The complex containing the Cas1 Y22A mutant regains marginal activity with substrates containing 5- or 6-nucleotide splayed ends, suggesting that Y22 steers the non-nucleophilic DNA strand away from the active site (Fig. 3d). Notably, the displaced non-nucleophilic strand is not cleaved into a shorter fragment by Cas1–Cas2, as the protospacer ends are not processed during integration (Extended Data Fig. 6c).

To determine the trajectory of the displaced non-nucleophilic strand after end-splaying, we crystallized Cas1–Cas2 with a protospacer with five-nucleotide frayed ends on both sides (Fig. 3a, b). The electron density at the fork is similar to the structures described above, except that we observe the first nucleotide of the displaced non-nucleophilic strand pointing in the opposite direction from the nucleophilic single-stranded DNA strand. Clear electron density is not observed for the remaining nucleotides of the displaced strand, indicating that they are not stabilized by the complex.

An alternative crystal form grown in the presence of Mg^{2+} reveals secondary Cas1–DNA interactions that provide additional insight into the mechanism of Cas1–Cas2 genomic DNA target binding and subsequent integration. In addition to the two Cas1 'catalytic' active sites carrying the 3′-OH ends of the protospacer, the 'non-catalytic' Cas1 active sites interact with the protospacer DNA from a symmetry mate, revealing a possible coordination of the target DNA during integration (Fig. 4a and Extended Data Fig. 7a). The non-catalytic Cas1 engages the DNA minor groove by contacts with α-helix 7, causing a slight kink on the DNA compared to our alternative crystal form lacking Mg^{2+} (Extended Data Fig. 7b). A close-up of the active site shows continuous density for Mg^{2+} with E141, H208, D221 and a phosphate backbone of the presumed target DNA, capturing a snapshot of scissile phosphodiester bond coordination before integration (Fig. 4a).

Because integration must occur in the active site that coordinates the 3′-OH of the protospacer DNA, we modelled the protein–DNA interactions from the non-catalytic Cas1 active sites into the catalytic Cas1 active sites. This reveals the positioning of the nucleophilic 3′-OH of the protospacer ends for attacking the scissile phosphodiester bond in the modelled DNA (Fig. 4b, c). Further work will be needed to shed light on how the complex specifically recognizes the leader-repeat region of the CRISPR locus for integration, as recently observed *in vitro*[11–13].

Together, these data explain key aspects of Cas1–Cas2 integrase-mediated acquisition of new DNA into bacterial genomes. First, we show that the substrates for integration are double-stranded DNA. Importantly, however, optimal substrates include a central 23 bp helical region flanked by five single-stranded nucleotides on each 3′ end. If substrates for CRISPR integration come from single-stranded DNA products of RecBCD, as recently suggested, they must somehow anneal or otherwise become double stranded before Cas1–Cas2 capture[20]. It remains unclear how the Cas1–Cas2 complex recognizes the AAG protospacer adjacent motif during protospacer selection, since the terminal nucleotides containing the 3′-OH nucleophiles are coordinated similarly in the Cas1 active sites (Fig. 1). Second, the Cas1–Cas2 integrase architecture specifies the precise length of integrated DNA, ensuring uniformity of spacer lengths within CRISPR loci. Finally, the structure-based model of DNA target sequence positioning suggests that in addition to catalysing the integration reaction, Cas1 plays a role in binding the target CRISPR locus. Target binding could possibly disrupt the structural symmetry observed in the crystal structure to coordinate the sequence-specific integration reactions at the leader-end of the CRISPR locus. Insights into target site recognition may offer strategies for altering or enhancing integration site specificity, with implications for use of the Cas1–Cas2 integrase as a genome-modifying technology.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 31 August; accepted 6 October 2015.
Published online 21 October 2015.

1. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315,** 1709–1712 (2007).
2. Mojica, F. J., Diez-Villasenor, C., Garcia-Martinez, J. & Soria, E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* **60,** 174–182 (2005).
3. Bolotin, A., Quinquis, B., Sorokin, A. & Ehrlich, S. D. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151,** 2551–2561 (2005).
4. Pourcel, C., Salvignol, G. & Vergnaud, G. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* **151,** 653–663 (2005).
5. Garneau, J. E. *et al.* The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468,** 67–71 (2010).
6. van der Oost, J., Westra, E. R., Jackson, R. N. & Wiedenheft, B. Unravelling the structural and mechanistic basis of CRISPR–Cas systems. *Nature Rev. Microbiol.* **12,** 479–492 (2014).
7. Yosef, I., Goren, M. G. & Qimron, U. Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Res.* **40,** 5569–5576 (2012).
8. Datsenko, K. A. *et al.* Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nature Comm.* **3,** 945 (2012).
9. Swarts, D. C., Mosterd, C., van Passel, M. W. & Brouns, S. J. CRISPR interference directs strand specific spacer acquisition. *PLoS ONE* **7,** e35888 (2012).
10. Nuñez, J. K. *et al.* Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR–Cas adaptive immunity. *Nature Struct. Mol. Biol.* **21,** 528–534 (2014).

11. Arslan, Z., Hermanns, V., Wurm, R., Wagner, R. & Pul, U. Detection and characterization of spacer integration intermediates in type I-E CRISPR–Cas system. *Nucleic Acids Res.* **42,** 7884–7893 (2014).
12. Nuñez, J. K., Lee, A. S., Engelman, A. & Doudna, J. A. Integrase-mediated spacer acquisition during CRISPR–Cas adaptive immunity. *Nature* **519,** 193–198 (2015).
13. Rollie, C., Schneider, S., Brinkmann, A. S., Bolt, E. L. & White, M. F. Intrinsic sequence specificity of the Cas1 integrase directs new spacer acquisition. *eLife* **4,** 10.7554/eLife.08716 (2015).
14. Heler, R., Marraffini, L. A. & Bikard, D. Adapting to new threats: the generation of memory by CRISPR–Cas immune systems. *Mol. Microbiol.* **93,** 1–9 (2014).
15. Brouns, S. J. *et al.* Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321,** 960–964 (2008).
16. Carte, J., Wang, R., Li, H., Terns, R. M. & Terns, M. P. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Dev.* **22,** 3489–3496 (2008).
17. Haurwitz, R. E., Jinek, M., Wiedenheft, B., Zhou, K. & Doudna, J. A. Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* **329,** 1355–1358 (2010).
18. Deltcheva, E. *et al.* CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471,** 602–607 (2011).
19. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337,** 816–821 (2012).
20. Levy, A. *et al.* CRISPR adaptation biases explain preference for acquisition of foreign DNA. *Nature* **520,** 505–510 (2015).
21. Wiedenheft, B. *et al.* Structural basis for DNase activity of a conserved protein implicted in CRISPR-mediated genome defense. *Structure* **17,** 904–912 (2009).
22. Savilahti, H., Rice, P. A. & Mizuuchi, K. The phage Mu transpososome core: DNA requirements for assembly and function. *EMBO J.* **14,** 4893–4903 (1995).
23. Scottoline, B. P., Chow, S., Ellison, V. & Brown, P. O. Disruption of the terminal base pairs of retroviral DNA during integration. *Genes Dev.* **11,** 371–382 (1997).
24. Katz, R. A., Merkel, G., Andrake, M. D., Roder, H. & Skalka, A. M. Retroviral integrases promote fraying of viral DNA ends. *J. Biol. Chem.* **286,** 25710–25718 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank G. Meigs and the 8.3.1 beamline staff at the Advanced Light Source for assistance with data collection, J. Chen for input on experimental design and members of the Doudna laboratory for comments and discussions. The 8.3.1 beamline is supported by UC Office of the President, Multicampus Research Programs and Initiatives grant MR-15-328599 and Program for Breakthrough Biomedical Research, which is partially funded by the Sandler Foundation. This project was funded by US National Science Foundation grant No. 1244557 to J.A.D. and by NIH grant AI070042 to A.N.E. J.K.N. and L.B.H. are supported by US National Science Foundation Graduate Research Fellowships and J.K.N. by a UC Berkeley Chancellor's Graduate Fellowship. P.J.K. is supported as a Howard Hughes Medical Institute Fellow of the Life Sciences Research Foundation. J.A.D. is an Investigator of the Howard Hughes Medical Institute and a member of the Center for RNA Systems Biology.

Author Contributions J.K.N. and L.B.H. conducted the crystallography, biochemistry and *in vivo* spacer acquisition assays. J.K.N., L.B.H. and P.J.K. collected the X-ray diffraction data and determined the crystal structures. J.K.N., L.B.H., P.J.K., A.N.E. and J.A.D. designed the study, analysed all data and wrote the manuscript.

Author Information Atomic coordinates and structure factors for the reported crystal structures have been deposited at the Protein Data Bank under accession codes 5DS4 (no Mg^{2+}), 5DS5 (with Mg^{2+}) and 5DS6 (splayed DNA). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.A.D. (doudna@berkeley.edu).

METHODS

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Cas1, Cas2 and DNA preparation. The Cas1 and Cas2 proteins from *E. coli* K12 (MG1655) were cloned and separately purified as previously described[10]. Single-stranded DNA (ssDNA) oligonucleotides purchased from Integrated DNA Technologies were annealed in 20 mM HEPES-NaOH, pH 7.5, 25 mM KCl, 10 mM $MgCl_2$ by heating at 95 °C for 3 min and slow cooling to room temperature. The pCRISPR DNA target for *in vitro* integration was constructed as previously described[12]. The DNA substrates used for crystallization were gel-purified before complex formation. The sequences for the five-nucleotide overhang substrates used for crystallization are: ssDNA1, 5′-ATTTACTACTCGTTCTGGTGTTTCT CGT-3′; and ssDNA2, 5′-AAACACCAGAACGAGTAGTAAATTGGGC-3′. The sequences for the five-nucleotide splayed substrates are: ssDNA1, 5′-TAAACAT TTACTACTCGTTCTGGTGTTTCTCGT-3′; and ssDNA2, 5′-CATCTAAACAC CAGAACGAGTAGTAAATTGGGC-3′.

***In vivo* acquisition and *in vitro* integration assays.** The *in vivo* acquisition assays were performed as previously described[7]. The *in vitro* integration reactions were conducted as previously described with slight modifications[12]. After pre-incubation of equimolar Cas1 and Cas2 at 4 °C, 100 nM of the resulting Cas1–Cas2 complex was incubated with 100 nM protospacer DNA for an additional 10–15 min at room temperature. The integration reaction was activated by the addition of 300 ng (~5 nM) pCRISPR, incubated at 37 °C for 1 h and quenched with DNA loading buffer supplemented with EDTA at a final concentration of 20 mM. The reaction products were analysed on 1.5% agarose gels. Per cent integration activity values were determined by quantifying the band intensity of the relaxed pCRISPR product and dividing over the intensity of all bands detected by Image Lab Software (Bio-Rad). We note that the integration activity could be a mixture of half-site and full-site integration products, as described previously[12].

Complex formation, crystallization and structure determination. Purified Cas1 and Cas2 were incubated with protospacer DNA at equimolar concentrations (50 μM) in buffer A (500 mM KCl, 20 mM HEPES-NaOH, pH 7.5, 1 mM DTT, 10 mM EDTA), followed by overnight dialysis at 4 °C against buffer B (100 mM KCl, 20 mM HEPES-NaOH, pH 7.5, 1 mM DTT, 5 mM EDTA). The dialysed sample was applied on a Superdex 75 10/300 column (GE Healthcare) in buffer B. Peak fractions were pooled and concentrated to ~3 mg ml^{-1} for crystallization. Optimized crystals were grown by hanging-drop vapour diffusion at room temperature in two different conditions, as described in the text. The Mg^{2+}-containing crystals grew as gem-like morphologies in 50 mM MES, pH 6.1, 10% isopropanol and 20 mM $MgCl_2$. The Mg^{2+}-free crystals grew as rods in 100 mM sodium citrate tribasic pH 5.6, 200 mM sodium acetate and 8% PEG 8000 (w/v). The crystals were briefly transferred into a drop containing either 25% ethylene glycol (with Mg^{2+} crystals) or 30% glycerol (without Mg^{2+} crystals) for cryoprotection and frozen in liquid nitrogen. The Cas1–Cas2 complex with a splayed DNA substrate crystallized in the same conditions as the Mg^{2+}-free crystals.

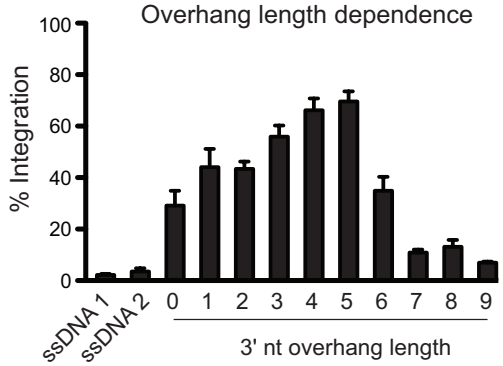
X-ray diffraction data were collected under cryogenic conditions at beamline 8.3.1 at the Lawrence Berkeley National Laboratory Advanced Light Source. Initial phases were obtained by sequential molecular replacement using individual protein components of the Cas1–Cas2 apo structure (Protein Data Bank (PDB) accession number 4P6I) as search models. Following initial placement of two Cas1 dimers and a dimer of Cas2, phases were improved by performing one round of rigid body refinement in PHENIX[25]. The resulting maps showed clear unbiased density for protospacer DNA, and subsequent model building was performed through iterative rounds of building in Coot[26] and refinement in PHENIX with NCS restraints on the protein subunits. The asymmetric unit of the three structures contains one copy of the Cas1–Cas2 complex bound to protospacer DNA. Statistics for the final crystal structures are reported in Extended Data Table 1. The final structures are missing clear density for the loop connecting α6 and α7 of Cas1. We assume this loop to be highly disordered as it is also not observed in the apo *E. coli* Cas1 crystal structure (PDB 3NKD) and the apo Cas1–Cas2 complex (PDB 4P6I)[10,27].

Fluorescence polarization. Fluorescence polarization assays were performed in 20 mM HEPES-NaOH, pH 7.5, 25 mM KCl, 5 mM EDTA, 1 μg ml^{-1} BSA and 1 mM DTT. Cas1–Cas2 were complexed and purified over gel filtration for all binding assays. The 3′-fluorescein labelled DNA substrate was added to the protein solution at a final concentration of 5 nM and the DNA–protein mixture was allowed to incubate for 30 min at 22 °C. Measurements were made by excitation at 485 nm and monitoring emission at 535 nm. Data were fit to a binding isotherm to obtain K_d. Each experiment was conducted in triplicate and error bars represent the standard deviation.

Sequence alignment. The *cas1* sequences were obtained from the National Center for Biotechnology Information (NCBI) Gene Data Bank. A representative *cas1* from each CRISPR type I subtype were chosen based on previous subtype assignments and the alignment was generated using MAFFT[28,29]. The organisms chosen for the alignment are: *Escherichia coli* K-12, *Cronobacter dublinensis* str. 582, *Erwinia amylovora*, *Yersinia pestis* biovar *Antiqua* str. B42003004, *Yersinia kristensenii*, *Hafnia alvei*, *Sulfolobus solfataricus*, *Thermotoga maritima*, *Pseudothermotoga lettingae*, *Deferribacter desulfuricans*, *Desulfovibrio vulgaris*, *Bacillus halodurans*, *Bacillus cereus*, *Synechocystis* sp. PCC 6803, *Cyanothece* sp. PCC 8802 and *Limnoraphis robusta*.

25. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66,** 213–221 (2010).
26. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60,** 2126–2132 (2004).
27. Babu, M. *et al.* A dual function of the CRISPR–Cas system in bacterial antivirus immunity and DNA repair. *Mol. Microbiol.* **79,** 484–502 (2011).
28. Makarova, K. S. *et al.* Evolution and classification of the CRISPR–Cas systems. *Nature Rev. Microbiol.* **9,** 467–477 (2011).
29. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30,** 772–780 (2013).

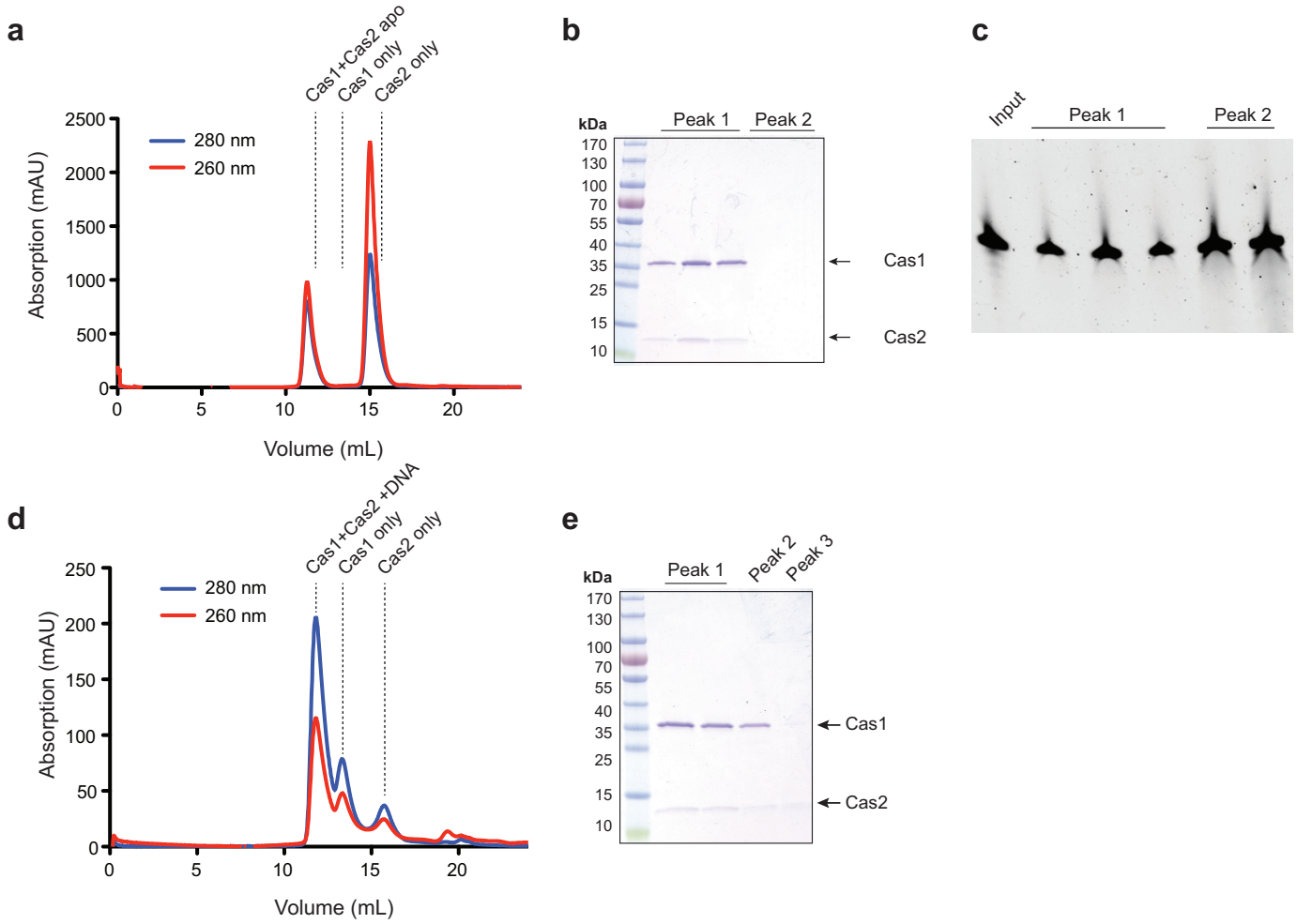
a



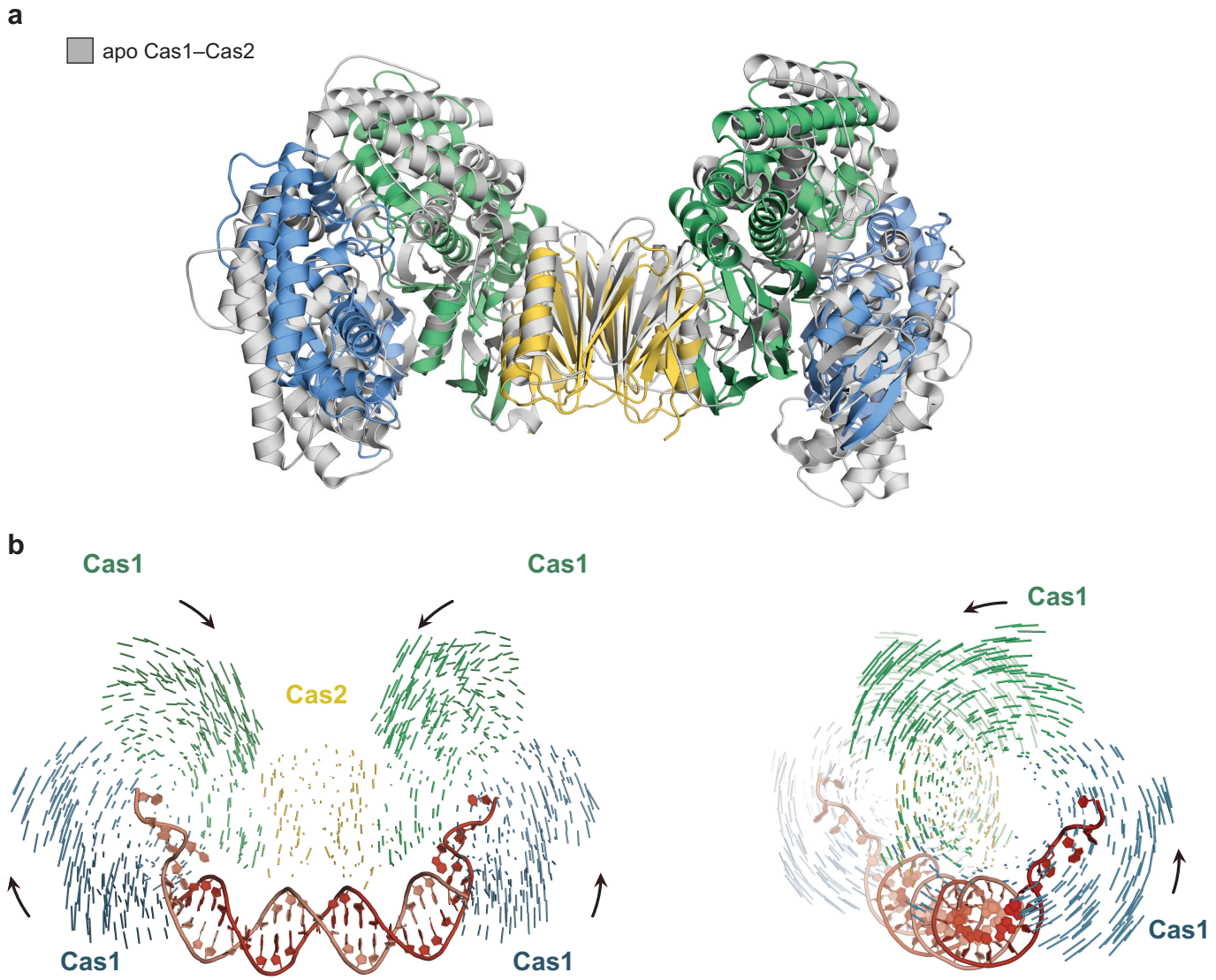
b

Full dsDNA	5' GCCCAATTTACTACTCGTTCTGGTGTTTCTCGT 3' 3' CGGGTTAAATGATGAGCAAGACCACAAAGAGCA 5'	5' ATTTACTACTCGTTCTGGTGTTTCTCGT 3' 3' CGGGTTAAATGATGAGCAAGACCACAAA 5' **5 nt**
1 nt	5' CCCAATTTACTACTCGTTCTGGTGTTTCTCGT 3' 3' CGGGTTAAATGATGAGCAAGACCACAAAGAGC 5'	5' TTTACTACTCGTTCTGGTGTTTCTCGT 3' 3' CGGGTTAAATGATGAGCAAGACCACAA 5' **6 nt**
2 nt	5' CCAATTTACTACTCGTTCTGGTGTTTCTCGT 3' 3' CGGGTTAAATGATGAGCAAGACCACAAAGAG 5'	5' TTACTACTCGTTCTGGTGTTTCTCGT 3' 3' CGGGTTAAATGATGAGCAAGACCACA 5' **7 nt**
3 nt	5' CAATTTACTACTCGTTCTGGTGTTTCTCGT 3' 3' CGGGTTAAATGATGAGCAAGACCACAAAGA 5'	5' TACTACTCGTTCTGGTGTTTCTCGT 3' 3' CGGGTTAAATGATGAGCAAGACCAC 5' **8 nt**
4 nt	5' AATTTACTACTCGTTCTGGTGTTTCTCGT 3' 3' CGGGTTAAATGATGAGCAAGACCACAAAG 5'	5' ACTACTCGTTCTGGTGTTTCTCGT 3' 3' CGGGTTAAATGATGAGCAAGACCA 5' **9 nt**

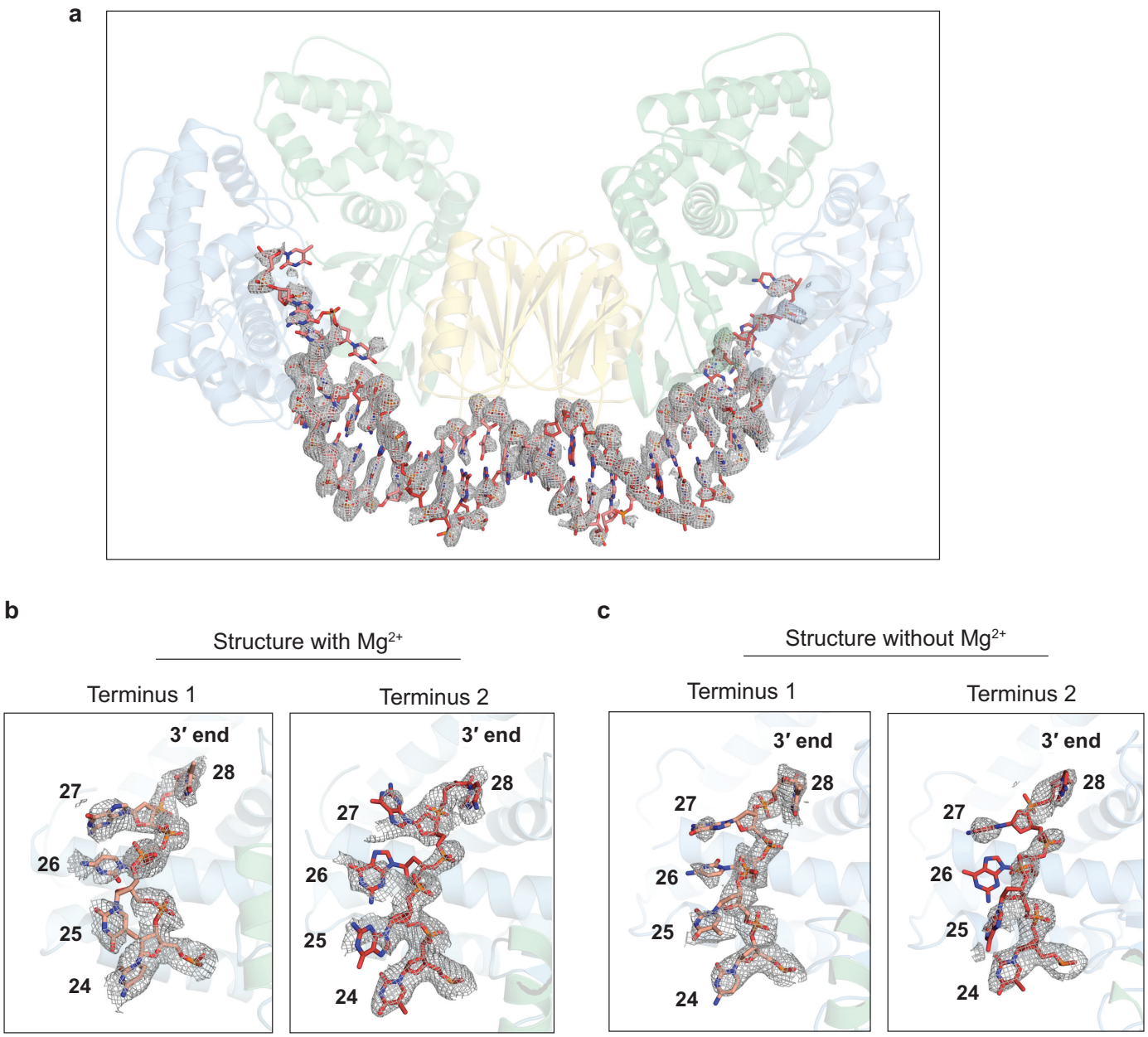
Extended Data Figure 1 | Effect of overhang length on integration efficiency. a, A plot of the per cent integration of protospacers ± standard deviation with varying 3′ single-stranded DNA extensions. A representative gel is shown in Fig. 1a. **b**, Protospacer sequences used for the assays described in **a** and Fig. 1a, with the red nucleotides indicating the 3′ overhang regions.



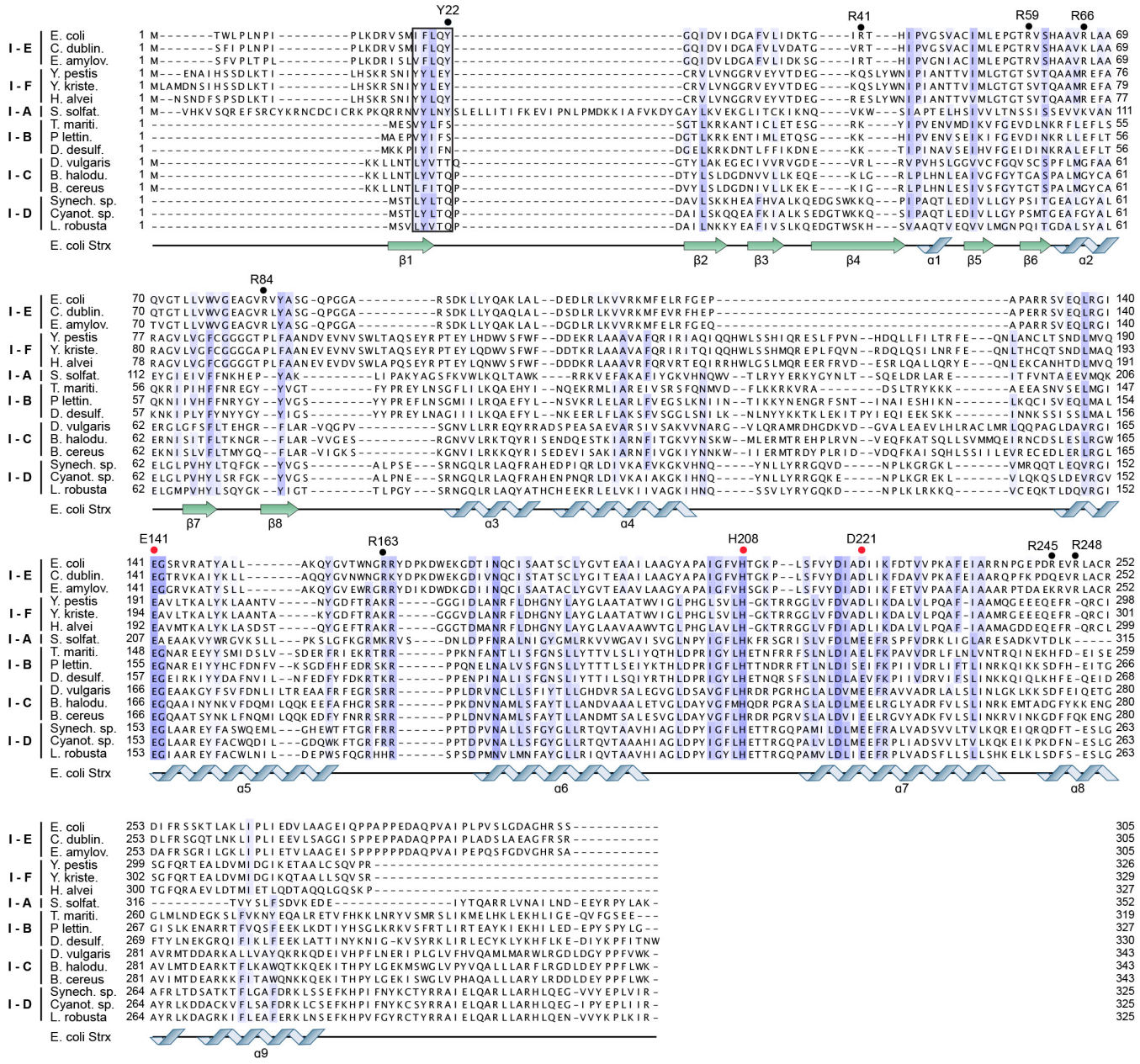
Extended Data Figure 2 | Assembly of Cas1–Cas2 complex bound to protospacer DNA. a, Gel filtration chromatogram of pre-assembled Cas1–Cas2 complex with protospacer DNA containing five-nucleotide 3′ overhangs. The dotted lines indicate the peak fractions of the Cas1–Cas2 complex without DNA, as shown in **d**. The dotted lines indicate the peak fractions of the Cas1–Cas2 complex bound to DNA (first peak) and excess, unbound DNA (second peak). **b**, **c**, The fractions from peak 1 (~12 ml) and peak 2 (~15 ml) were analysed by Coomassie-stained SDS–PAGE (**b**) and 12% urea-PAGE (**c**) to confirm the presence of Cas1, Cas2 and protospacer DNA. **d**, Gel-filtration chromatogram of assembled Cas1–Cas2 without protospacer DNA. **e**, Coomassie-stained SDS–PAGE of the peak fractions from **d**. Supplementary Information contains the full images for **b**, **c** and **e**.



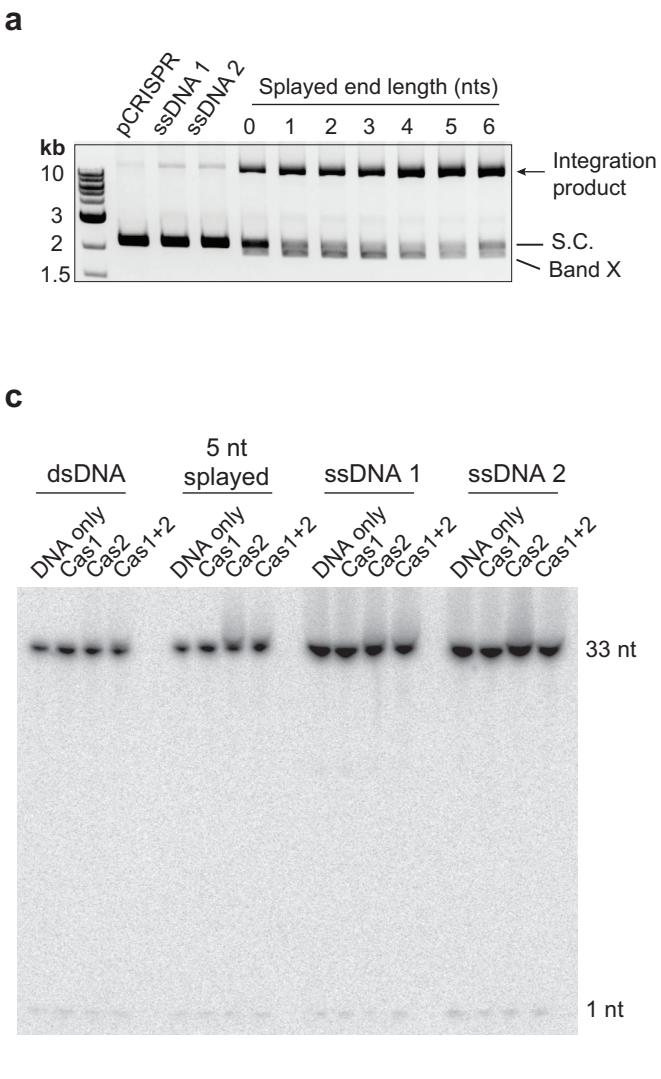
Extended Data Figure 3 | Conformational dynamics upon protospacer DNA binding. a, An overlay of the DNA-bound Cas1–Cas2 structure with the apo Cas1–Cas2 (grey, PDB 4P6I). **b**, Vector lines depicting the conformational changes the Cas1–Cas2 complex undergoes upon protospacer DNA binding compared to the apo complex (PDB 4P6I). The Cas1 subunits rotate towards the direction of the arrows.



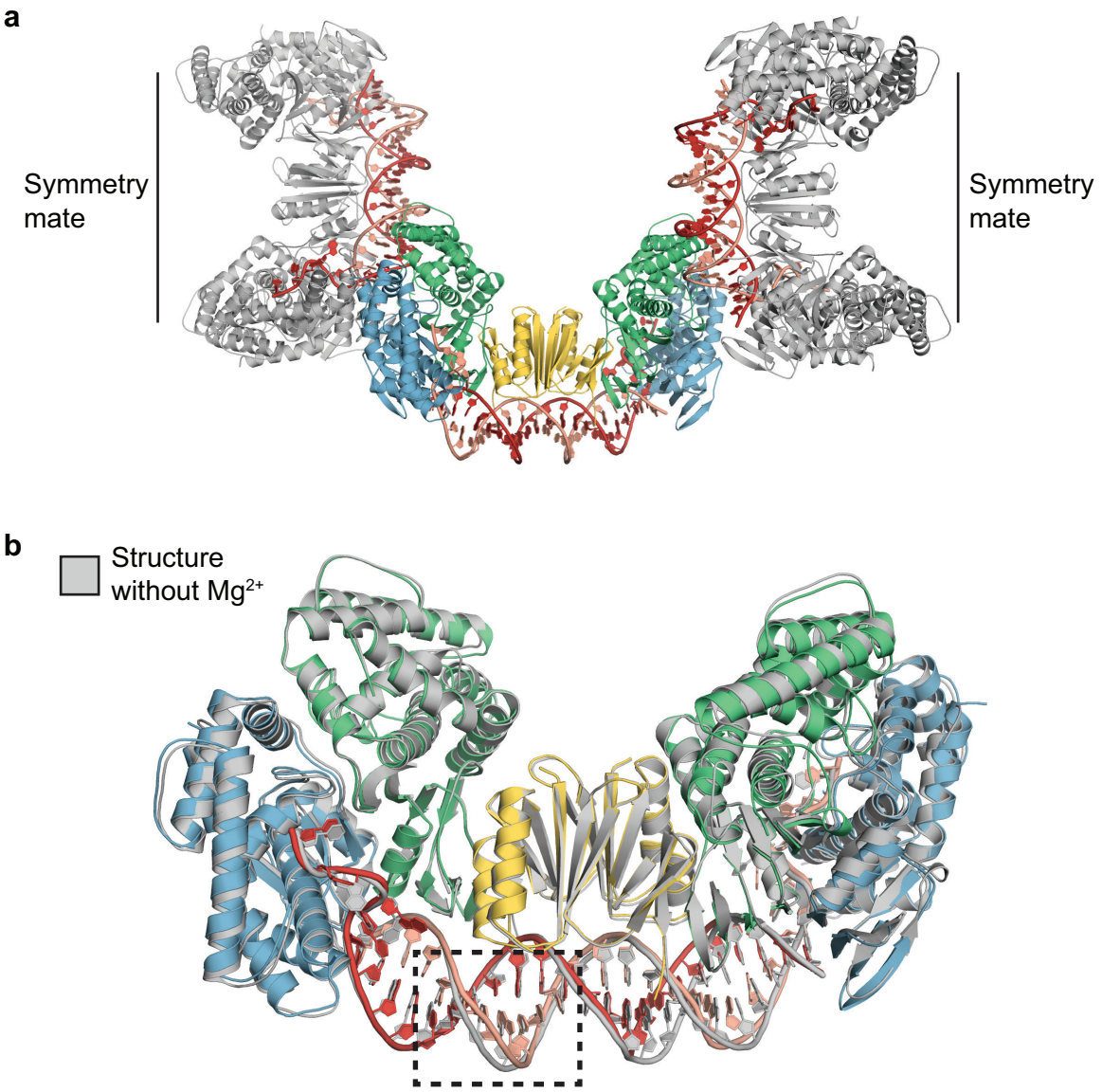
Extended Data Figure 4 | Omit maps of the protospacer DNA. a, Simulated annealing $F_o - F_c$ omit electron density map of the entire protospacer DNA using the 'no Mg^{2+}' map and model. **b**, **c**, Simulated annealing $F_o - F_c$ omit electron density maps of the terminal five nucleotides in the active sites of the structures (**a**) with Mg^{2+} or (**b**) without Mg^{2+} in the crystallization condition. The maps are contoured at 2.0σ.



Extended Data Figure 5 | Sequence alignment of Cas1 proteins in type I CRISPR systems. Sequence alignments of Cas1 from representative organisms with type I CRISPR systems. The *E. coli* sequence is displayed at the top. The dots indicate the residues described in this study, with the red dots indicating the metal-binding residues. The box highlights the non-universal conservation of the *E. coli* Y22 residue in the β1 region of type I CRISPR systems. The secondary structure representations shown are for the *E. coli* Cas1.



Extended Data Figure 6 | Integration of protospacer substrates with splayed ends. a, Representative agarose gel of *in vitro* integration reactions using increasing lengths of splayed ends. The average per cent integration of three independent experiments is plotted in Fig. 3d. **b**, Sequences of protospacers used in the integration assays in **a**. **c**, A 12% denaturing polyacrylamide gel of protospacers after incubation with Cas1–Cas2 for 1 h at 37 °C in integration assay buffer conditions. The indicated DNA substrates are radiolabelled at the 5′ end. Supplementary Information contains the full images for **a** and **c**. nt, nucleotide.



Extended Data Figure 7 | Crystallographic packing of the complex bound to Mg^{2+}. a, View of the symmetry mates (grey) contacting the non-catalytic Cas1 subunits (green). Catalytic Cas1 subunits are shown in blue, Cas2 in yellow and DNA is shown in salmon and red. **b**, Superposition of our two crystal structures, with or without Mg^{2+}, shows a slight DNA kink in the structure bound to Mg^{2+} (dotted box). This region contacts α-helix 7 of a symmetry mate, as described in the text.

Extended Data Table 1 | Summary of X-ray crystallography data collection and refinement

	Without Mg^{2+}	With Mg^{2+}	Splayed substrate
Data collection			
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions			
a, b, c (Å)	88.02, 120.01, 196.01	75.66, 165.93, 167.26	88.02, 123.01, 196.01
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	49.00–3.20 (3.36 – 3.20)	46.41–2.95 (3.06–2.95)	48.9–3.35 (3.42–3.35)
R_{merge} (%)	30.8 (146)	19.6 (157)	28.5 (126)
R_{pim} (%)	12.8 (61.4)	10.8 (86.3)	21.6 (94.3)
I/σ	6.4 (1.5)	9.8 (1.4)	5.0 (1.3)
$CC_{1/2}$	98.5 (72.4)	99.3 (42.0)	98.3 (72.7)
Completeness (%)	99.8 (99.0)	100 (99.9)	99.6 (97.7)
Redundancy	6.7 (6.6)	7.9 (8.0)	4.1 (4.0)
Wilson B factor ($Å^2$)	63.8	64.0	73.7
Refinement			
Resolution (Å)	49.00–3.20	46.41–2.95	49.00–3.35
No. reflections	35,808 (3,502)	44,960 (4,418)	31,049 (2885)
R_{work}/R_{free}	24.2/27.0	23.0/25.4	23.2/27.4
No. atoms			
Protein	9,375	9,576	9,375
DNA	1,142	1,142	1,165
Metal	0	4	0
Average B-factors ($Å^2$)			
Protein	65.9	66.6	86.6
DNA	76.2	67.2	103.0
Metal		51.6	
R.m.s deviations			
Bond lengths (Å)	0.003	0.003	0.004
Bond angles (°)	0.72	0.75	0.81
Ramachandran statistics (%)			
Favored	96.0	95.0	96.0
Allowed	3.75	4.51	3.58
Outliers	0.25	0.49	0.42

One crystal was used for each structure.
Highest resolution shell is shown in parenthesis.

LETTER

doi:10.1038/nature15519

Endoperoxide formation by an α-ketoglutarate-dependent mononuclear non-haem iron enzyme

Wupeng Yan[1]*, Heng Song[2]*, Fuhang Song[3], Yisong Guo[6], Cheng-Hsuan Wu[2], Ampon Sae Her[2], Yi Pu[2,4], Shu Wang[2], Nathchar Naowarojna[2], Andrew Weitz[6], Michael P. Hendrich[6], Catherine E. Costello[2,4], Lixin Zhang[3], Pinghua Liu[2] & Yan Jessie Zhang[1,5]

Many peroxy-containing secondary metabolites[1,2] have been isolated and shown to provide beneficial effects to human health[3–5]. Yet, the mechanisms of most endoperoxide biosyntheses are not well understood. Although endoperoxides have been suggested as key reaction intermediates in several cases[6–8], the only well-characterized endoperoxide biosynthetic enzyme is prostaglandin H synthase, a haem-containing enzyme[9]. Fumitremorgin B endoperoxidase (FtmOx1) from *Aspergillus fumigatus* is the first reported α-ketoglutarate-dependent mononuclear non-haem iron enzyme that can catalyse an endoperoxide formation reaction[10–12]. To elucidate the mechanistic details for this unique chemical transformation, we report the X-ray crystal structures of FtmOx1 and the binary complexes it forms with either the co-substrate (α-ketoglutarate) or the substrate (fumitremorgin B). Uniquely, after α-ketoglutarate has bound to the mononuclear iron centre in a bidentate fashion, the remaining open site for oxygen binding and activation is shielded from the substrate or the solvent by a tyrosine residue (Y224). Upon replacing Y224 with alanine or phenylalanine, the FtmOx1 catalysis diverts from endoperoxide formation to the more commonly observed hydroxylation. Subsequent characterizations by a combination of stopped-flow optical absorption spectroscopy and freeze-quench electron paramagnetic resonance spectroscopy support the presence of transient radical species in FtmOx1 catalysis. Our results help to unravel the novel mechanism for this endoperoxide formation reaction.

The verruculogen biosynthetic gene cluster was identified through bioinformatic analysis[10], and its chemical scaffold is assembled by a non-ribosomal peptide synthetase, followed by several tailoring reactions. Among them, the FtmOx1-catalysed endoperoxide formation reaction is the most notable (Fig. 1a). Recent biochemical characterizations indicate that, unlike prostaglandin H synthase, FtmOx1 is an α-ketoglutarate (α-KG)-dependent mononuclear non-haem iron enzyme[11,12]. Further characterization indicates that molecular oxygen (O_2) is incorporated into verruculogen without O–O bond scission[11,12], distinguishing FtmOx1 from all currently known α-KG-dependent mononuclear non-haem iron enzymes[13–17].

To unravel the mechanistic details of FtmOx1 catalysis, we first characterized the FtmOx1–α-KG complex using anaerobically purified and Fe^{2+}-reconstituted FtmOx1 (FtmOx1-Fe^{II}). Upon mixing the reconstituted enzyme with α-KG under anaerobic conditions, a pink species appeared (pink trace, extinction coefficient ε_{520} of $\sim$166 M^{-1} cm^{-1}, Fig. 1b). The dissociation constant (K_d) of this species was $\sim 185 \pm 35$ μM (Extended Data Fig. 1a), close to the K_d values of Fe^{II}–α-KG complexes of other mononuclear non-haem iron enzymes (for example, that of TauD)[18]. Upon exposure to O_2 and in the absence of the substrate fumitremorgin B (**1**), the pink species faded and a blue chromophore with a λ_{max} at $\sim$600 nm developed within 30 min (blue trace, Fig. 1b). Tandem mass spectrometry (MS/MS) analysis of the blue species indicated the oxidation of Y224 to dihydroxyphenylalanine (DOPA, Fig. 1c), which is the result of a self-hydroxylation reaction as observed in other mononuclear non-haem iron enzymes[19]. Notably, the presence of the substrate fumitremorgin B (**1**) prevented the FtmOx1 self-hydroxylation reaction (Extended Data Fig. 2). All of these properties are consistent with the formation of the FtmOx1-Fe^{II}–α-KG complex.

In previous studies, ascorbate was included as an additional reductant[11,12], although its role in FtmOx1 catalysis was not known[11,12]. We observed that FtmOx1 is capable of catalysing fumitremorgin B (**1**) oxidation in the absence of ascorbate (Fig. 1d and Extended Data Fig. 3). At a fixed FtmOx1-Fe^{II}:fumitremorgin-B ratio of 1:1.5 and with an excess of O_2, the amount of product increased with the amount of α-KG until the α-KG:FtmOx1-Fe^{II} ratio reached 1.0 (Fig. 1e and Extended Data Fig. 4). In contrast, when the O_2:FtmOx1-Fe^{II} ratio was below 1.0, only a small amount ($\sim$0.2 equivalent) of product was formed. Above 1.0, the amount of product increased with the increasing amount of O_2, and plateaued when the O_2:FtmOx1-Fe^{II} ratio was $>$2.0 (Fig. 1f and Extended Data Fig. 4). These results strongly suggest that each FtmOx1-catalysed turnover consumes one equivalent of α-KG and two equivalents of O_2. Unexpectedly, under our assay conditions, compound **3** rather than verruculogen (**2**) was the dominant product (Fig. 1e, f and Supplementary Information).

We determined the FtmOx1 crystal structure at 1.95 Å resolution with the phase derived from selenomethionine-labelled FtmOx1 using the single-wavelength anomalous dispersion method. FtmOx1 folds as a 'jelly roll', a prevalent fold in mononuclear non-haem iron enzymes (Fig. 2a)[13]. Two molecules in each asymmetric unit form a functional dimer, consistent with our size-exclusion chromatography profile and previous literature reports[11,12]. The dimer interface (2,461.6 $Å^2$) accounts for 17.1% of the FtmOx1 surface. The active site pocket at the dimer interface has a volume of 222.6 $Å^3$, as calculated by the DogSite Server[20]. This spacious pocket is partitioned into two parts: a hydrophilic region where the non-haem iron centre is located and a hydrophobic pocket formed by L64, F115, and F233 from one monomer with I267 and V268 from the other (Fig. 2b).

H129, H205, D131, and three well-ordered water molecules form an approximate octahedral coordination to the mononuclear iron (Fig. 2c). One of the water ligands is hydrogen-bonded to Y224, whose proximity to the mononuclear iron centre ($\sim$4.4 Å) explains the formation of DOPA in the FtmOx1 self-hydroxylation reaction (Fig. 1c). Co-crystallization or soaking of the co-substrate α-KG led to an identical FtmOx1-Fe^{II}–α-KG complex, in which α-KG binds to the iron centre in a bidentate fashion by replacing two water molecules

[1]Department of Molecular Biosciences, University of Texas at Austin, Austin, Texas 78712, USA. [2]Department of Chemistry, Boston University, Boston, Massachusetts 02215, USA. [3]CAS Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China. [4]Center for Biomedical Mass Spectrometry, Boston University School of Medicine, Boston, Massachusetts 02118, USA. [5]Institute for Cellular and Molecular Biology, University of Texas at Austin, Austin, Texas 78712, USA. [6]Department of Chemistry, Carnegie Mellon University, 4400 Fifth Avenue, Pittsburgh, Pennsylvania 15213, USA.
*These authors contributed equally to this work.

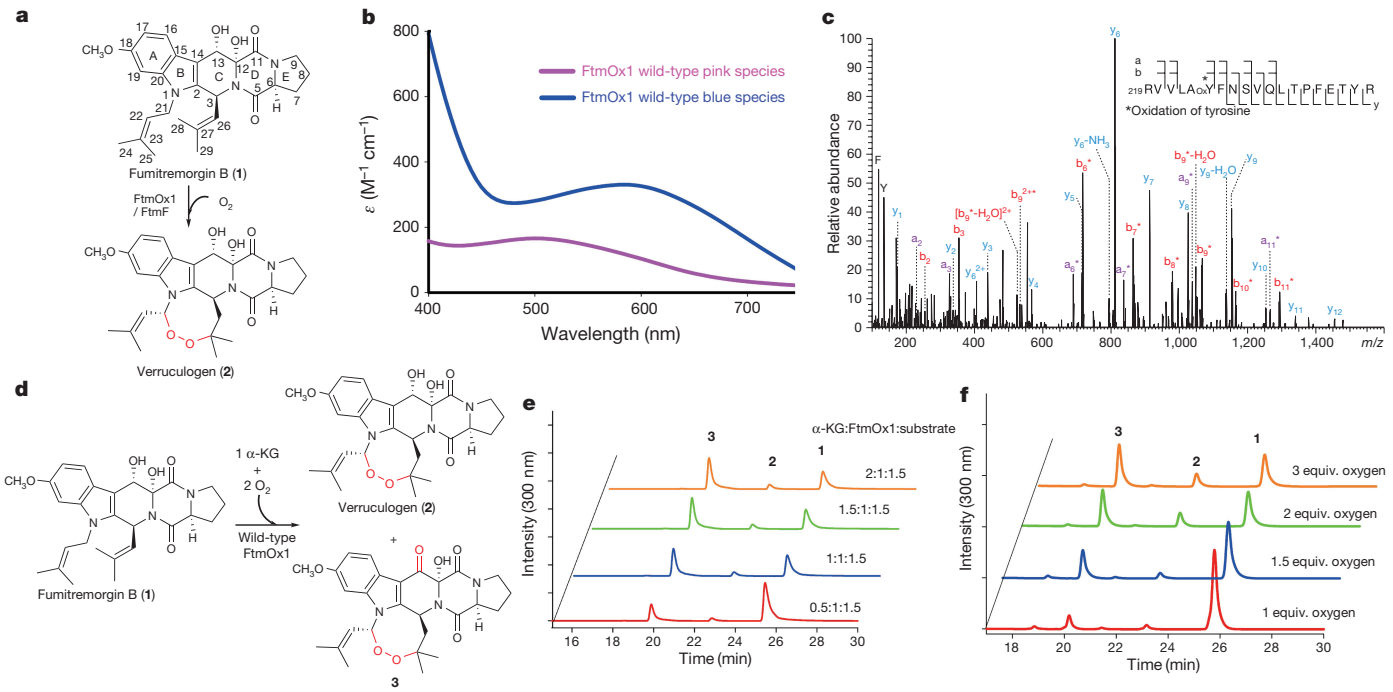


Figure 1 | Enzymatic characterization of wild-type FtmOx1. a, Proposed FtmOx1 reaction. **b**, Formation of the FtmOx1-Fe^{II}–α-KG binary complex under anaerobic conditions (pink trace). Self-hydroxylation reaction upon exposure of the binary complex to O_2 (blue trace). **c**, Electrospray ionization MS/MS analysis of the blue species in **b** is consistent with the oxidation of Y224 to DOPA224. **d**, Products formed in the FtmOx1 reaction. **e**, α-KG stoichiometry analysis in FtmOx1 catalysis. High-performance liquid chromatography (HPLC) chromatograms of FtmOx1 reactions with various amounts of α-KG. Identities of the peaks were assigned based on nuclear magnetic resonance (NMR) and high-resolution mass spectrometry (see Supplementary Information). **f**, O_2 stoichiometry analysis in FtmOx1 catalysis. HPLC chromatograms of FtmOx1 reactions contained fumitremorgin B (360 μM), FtmOx1 (240 μM), and α-KG (480 μM) when variable amounts of oxygen-saturated buffer were added to initiate the reaction.

(Fig. 2d). In the FtmOx1-Fe^{II}–α-KG complex, the 2-keto group of α-KG coordinates to the iron centre *trans* to D131. Its 1-carboxylate group binds *trans* to H205, which is the distal histidine of the 2-His-1-carboxylate facial triad (Fig. 2d and Extended Data Fig. 5a)[13]. In this FtmOx1-Fe^{II}–α-KG complex, the remaining water ligand (a potential site for O_2 binding and activation) is completely shielded from solvent or substrate by Y224 (Fig. 2d and Supplementary Video 1). In contrast, in most reported structures of enzyme–α-KG complexes, the 1-carboxylate of α-KG coordinates *trans* to the proximal histidine of the facial triad motif[13], and the remaining open site for O_2 binding and activation directly points towards the substrate. As a result, the oxoferryl (Fe^{IV}=O) species produced from oxygen activation is accessible to the substrate for oxidative transformations (for example, TauD in Fig. 2e and Supplementary Video 2)[21].

To examine whether Y224 changes location upon substrate binding, we also solved the structure of the FtmOx1-Fe^{II}–fumitremorgin-B complex at a resolution of 2.2 Å (Fig. 2f). The location of the positive density is consistent for all data sets collected for this complex by either co-crystallization or soaking (>15 data sets). Substrate is modelled into the density at the active site with an average occupancy of ~60% owing to the high hydrophobicity of fumitremorgin B (Extended Data Fig. 5b). In this complex, Y224 adopts a conformation identical to that observed in the FtmOx1 alone (Fig. 2c) and FtmOx1-Fe^{II}–α-KG complexes (Fig. 2d). Rings A and B of the substrate form π–π stacking with the Y224 side chain at a distance of ~3.3 Å (Fig. 2f). Superimposition of the structures of the FtmOx1-Fe^{II}–α-KG complex onto the FtmOx1-Fe^{II}–fumitremorgin-B complex revealed that the side chain of Y224 effectively separates the potential O_2 binding site from the substrate binding pocket (Fig. 2g, Extended Data Fig. 5c and Supplementary Video 1). This is in notable contrast to TauD, in which the oxygen binding and activation site directly faces the substrate (Extended Data Fig. 5d and Supplementary Video 2)[21].

On the basis of the strategic positioning of Y224 in the FtmOx1 active site, we next examined the role it plays in FtmOx1 catalysis. We characterized two Y224 variants, Y224A- and Y224F-substituted FtmOx1. The enzyme–α-KG complexes of both variants exhibit K_d values close to that of the wild-type FtmOx1 (Extended Data Fig. 1b, c). However, the product profiles of both variants were very different (Fig. 3). The Y224A-substituted variant produced a mixture of at least five detectable products with mainly dealkylation products (compounds **4**, **5** and **6**). Endoperoxides (**2** and **3**) only account for ~15% of the product population (Fig. 3a, b).

The Y224F-substituted variant also produced endoperoxides (**2** and **3**) and dealkylation products (**4** and **5**). In addition, there were more endoperoxides (**2** and **3**) formed by the FtmOx1(Y224F) variant relative to the FtmOx1(Y224A) variant (~35% versus ~15% of the product mixture, Fig. 3a). For the FtmOx1(Y224F)-Fe^{II}–α-KG complex in the absence of the substrate fumitremorgin B (**1**), exposure to O_2 caused the complex to slowly change colour to blue, which implies DOPA formation (Extended Data Fig. 6a). DOPA formation can be explained by two sequential hydroxylation steps (F224→Y224 and Y224→DOPA224). Indeed, this conclusion was supported by MS/MS analysis of this variant (Extended Data Fig. 6b–e). Thus, the higher level of endoperoxides (**2** and **3**) produced by the FtmOx1(Y224F) variant is probably attributed to the conversion of the variant to wild-type FtmOx1, which provides further evidence supporting the key role of Y224 in FtmOx1 catalysis.

Mononuclear non-haem iron enzymes catalyse a wide range of reactions[13–17]. Recently, unique transformations have been reported which demonstrate the functional versatility of this class of enzymes, including oxidative dehydrogenation in epoxide formation[22], chlorination[23,24], epimerization[25,26], and C–C bond cleavage[27]. FtmOx1 provides a further example of this diversity[11,12]. On the basis of our structural and biochemical information, we propose a preliminary FtmOx1 mechanistic model (Fig. 4). After α-KG and substrate binding,

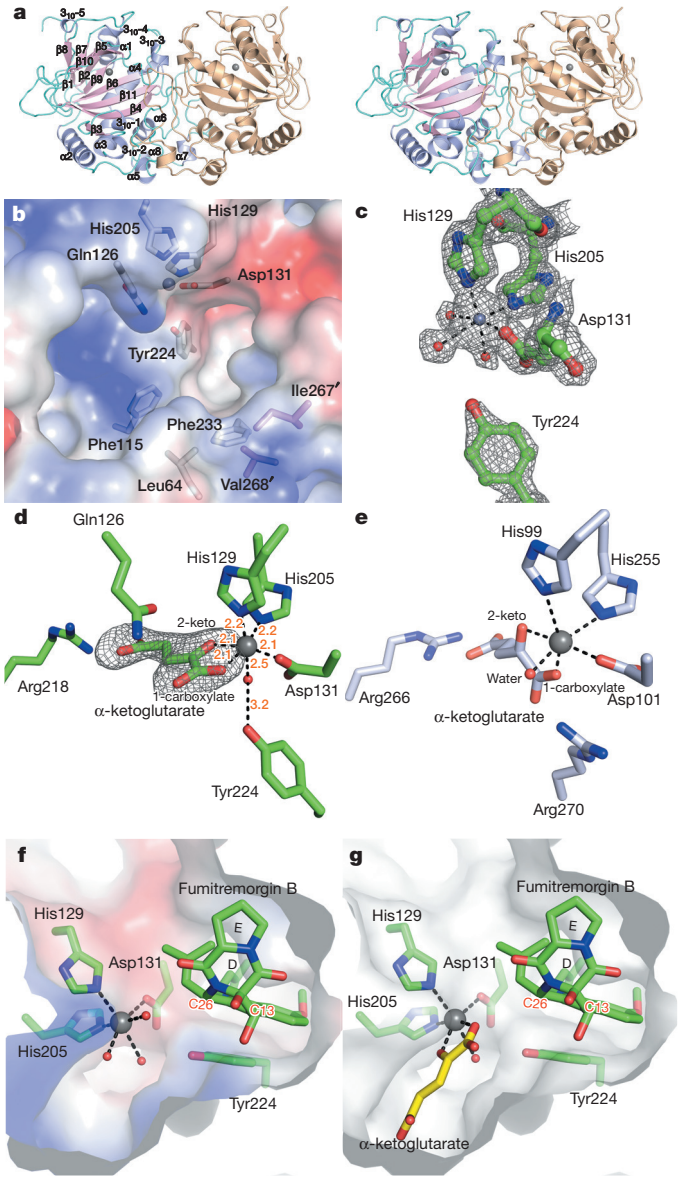


Figure 2 | Structures of FtmOx1. a, Overall architecture of FtmOx1 shown as a functional dimer with one monomer colour-coded based on secondary structures (shown as stereo images). The iron centre is labelled as a grey sphere. **b**, FtmOx1 active site shown in the electrostatistic mode. **c**, FtmOx1 metallo-centre electron density ($2mF_o - DF_c$ map) at 1σ contour. The coordination of iron is represented by dashed lines. **d**, FtmOx1-Fe^{II}–α-KG binary complex. The α-KG molecule was modelled into a composite omit map ($mF_o - DF_c$ map) contoured to 2.8σ. The coordination of iron is represented by dashed lines with distances labelled (units, Å). **e**, α-KG binding mode of TauD (PBD accession code 1OS7). TauD is shown in an identical orientation relative to that of FtmOx1 in **d** to highlight their differences in active site topologies. **f**, Structure of the FtmOx1-Fe^{II}–fumitremorgin-B complex. **g**, Superimposition of the binary structures of FtmOx1-Fe^{II}–α-KG and FtmOx1–fumitremorgin-B. Y224 is highlighted in pink.

the first molecule of O_2 is activated to produce an Fe^{IV}=O species (species **B**) by a mechanism similar to that of other members of this class of enzymes[28]. Uniquely, in FtmOx1, because the O_2 activation site is shielded from the substrate by Y224, direct oxidation of the substrate by the Fe^{IV}=O species is less likely. Instead, the Fe^{IV}=O species oxidizes Y224 to a tyrosyl radical (species **C**), which then removes a hydrogen atom from the fumitremorgin B C21 position to form a substrate-based radical (species **D**). A second molecule of O_2 reacts with species **D** to form a peroxyl radical (species **E**). It then reacts with the other prenyl arm to produce the endoperoxide along with the formation of a carbon

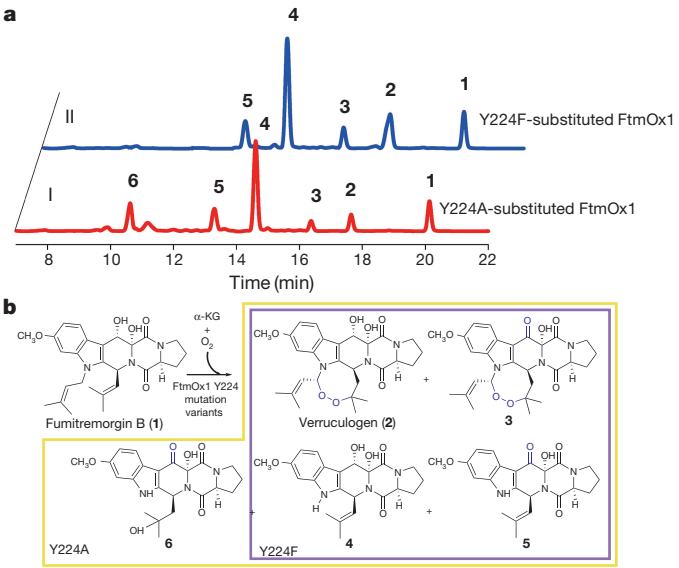


Figure 3 | Characterization of Y224A- and Y224F-substituted FtmOx1. a, HPLC profiles of reactions from the two FtmOx1 variants, Y224A and Y224F. Both traces were conducted in reaction mixture containing FtmOx1 Y224 variants (240 μM), fumitremorgin B (240 μM), α-KG (720 μM) and O_2 (480 μM). Trace I, HPLC chromatograms of the reaction using Y224A-substituted FtmOx1; trace II, HPLC chromatograms of the reaction using Y224F-substituted FtmOx1. Note that a new column was used for the mutant analyses relative to the one in wild-type FtmOx1 characterizations, which led to the differences in retention times relative to the other HPLC traces. **b**, Products formed in reactions using either Y224F- or Y224A-substituted FtmOx1 variants. The compounds were characterized using NMR and MS (see Supplementary Information).

centre radical at the C26 position (species **F**). Species **F** can re-oxidize Y224 to a tyrosyl radical (species **G**). Starting from species **G**, two pathways are possible. FtmOx1 can follow a mechanism similar to prostaglandin synthase H (ref. 9) in which once the tyrosyl radical is formed, multiple cycles of endoperoxide formation can be mediated through this radical (pathway I, Fig. 4). However, the production of compound **3** in the FtmOx1 reaction points to another possibility (pathway II, Fig. 4), in which the two electrons provided by the **2**→**3** oxidation process reduce both Fe^{3+} and the tyrosyl radical to the resting state of FtmOx1 (species **A**).

The formation of a small amount of endoperoxides in Y224A- and Y224F-substituted FtmOx1 may be due to two competing pathways: the hydroxyl-rebound and the endoperoxide formation pathways (Extended Data Fig. 7). After Fe^{IV}=O is formed, it may directly remove a hydrogen atom from the fumitremorgin B (**1**) C21 position to form a substrate-based radical (species **C′**, Extended Data Fig. 7). Subsequent rebound by the hydroxyl radical will lead to the formation of hydroxylation products (pathway I′, Extended Data Fig. 7). Decomposition of the hydroxylation reaction product forms compounds **4** and **5**. At the same time, the substrate-based radical (species **C′**) may be trapped by a second molecule of O_2, which leads to endoperoxide formation (pathway II′, Extended Data Fig. 7).

To gain evidence supporting the presence of radical species in FtmOx1 catalysis as outlined in our FtmOx1 mechanistic model (Fig. 4), we conducted spin-trapping experiments using 5,5-dimethyl-1-pyrroline *n*-oxide (DMPO) as the reagent. In the presence of 50 equivalents of DMPO, further oxidation of verruculogen **2** to **3** was markedly suppressed, and **2** was the dominant product (Fig. 5a). This result provides evidence supporting the involvement of radicals in FtmOx1 catalysis. Next, the FtmOx1 reaction was monitored with stopped-flow optical absorption spectroscopy (Fig. 5b). The UV–visible spectrum of the solution generated after rapid mixing of O_2-saturated buffer with the FtmOx1-Fe^{II}–fumitremorgin-B–α-KG complex demonstrated the accumulation of a transient species centred at ~420 nm. The amount

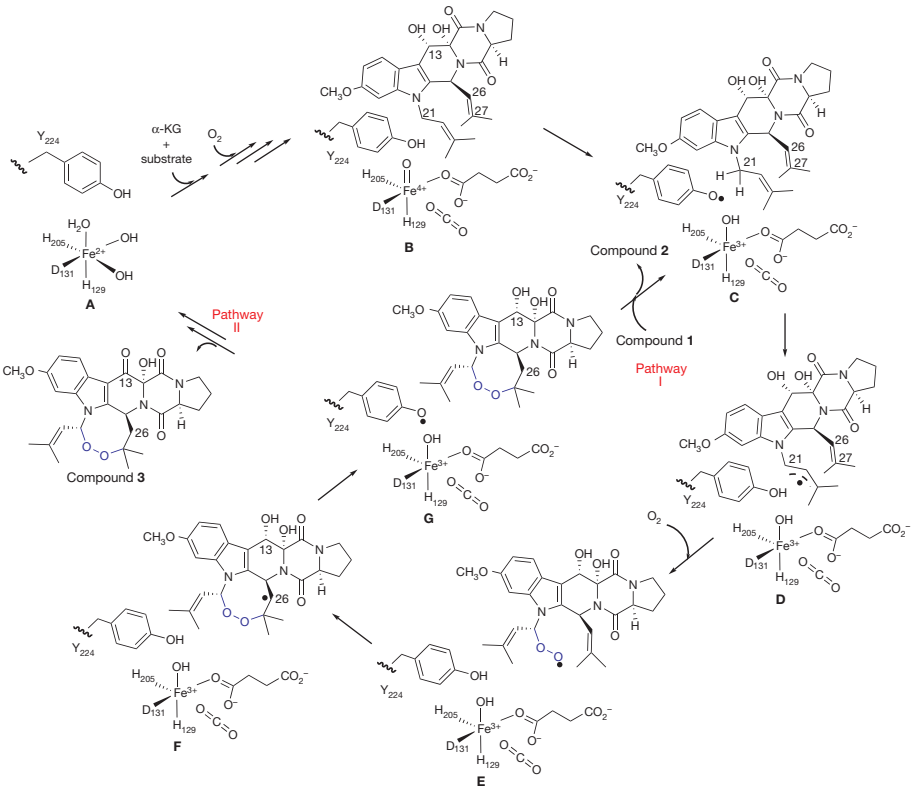


Figure 4 | Proposed FtmOx1 mechanistic model. The oxygen–oxygen bonds shown in blue highlight the incorporation of endoperoxide into the substrate fumitremorgin B.

of this species maximized at ~0.2 s and then decayed within ~3 s. This wavelength differs from the tyrosyl radicals observed in ribonucleotide reductase[29] and another reported α-KG-dependent iron enzyme, CarC[26] (a peak at 410 nm with a shoulder at 390 nm). Chemical quench experiments performed under the same conditions indicated that the consumption of substrate **1** and the formation of products (**2** and **3**) occurred on the timescale of seconds per cycle (Extended Data Fig. 8a), suggesting that the 420 nm species observed in the stopped-flow optical absorption spectroscopy experiments is a kinetically competent intermediate. FtmOx1 catalysis was then investigated by rapid freeze-quench in conjunction with electron paramagnetic resonance (EPR) spectroscopy. Two EPR signals were observed at 0.01 s (earliest possible time on instrument) and were highest at ~0.2 s after the rapid mixing of O_2-saturated buffer with the FtmOx1-Fe^{II}–fumitremorgin-B–α-KG complex (Extended Data Fig. 9a). The first EPR signal with resonances at $g = 4.54$, 4.26, and 3.93 (Extended Data Fig. 9b) belongs to a high-spin Fe^{3+} species having axial and rhombic zero-field splitting parameters of $|D| < 0.5\ cm^{-1}$ and $E/D \approx 0.26$, respectively. These parameters are not typical of adventitious Fe^{3+}. The second EPR signal was in the $g = 2$ region and most likely belongs to a radical species (Fig. 5c

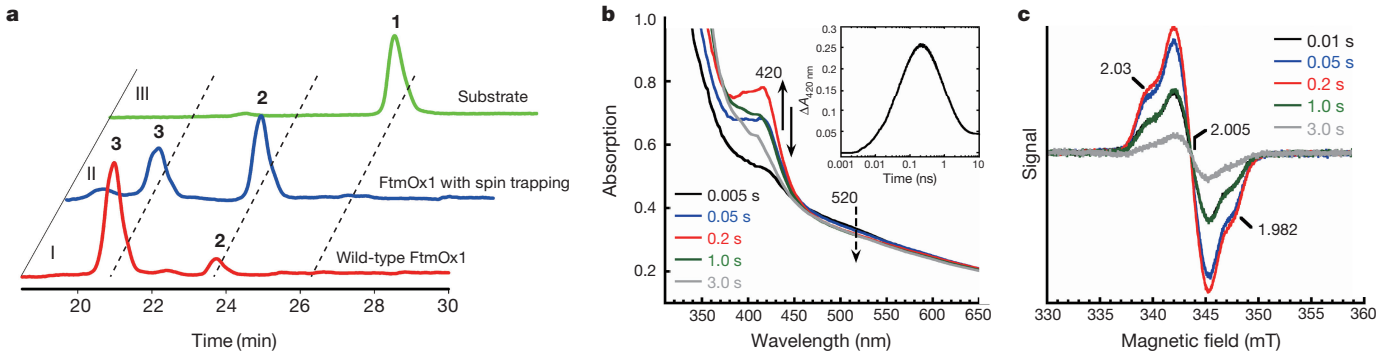


Figure 5 | Evidence for transient radical species in the reaction pathway. **a**, HPLC chromatograms of FtmOx1 reaction under three different conditions. The reaction mixture contained FtmOx1 (240 μM), fumitremorgin B (200 μM), α-KG (300 μM), and was initiated with O_2-saturated buffer. Trace I, FtmOx1 reaction; trace II, FtmOx1 reaction in the presence of 10 mM DMPO; trace III, FtmOx1 substrate alone. **b**, Absorbance changes upon mixing the O_2-saturated buffer with the reaction mixture in 100 mM Tris-HCl (pH 7.5) buffer containing FtmOx1 (0.65 mM), Fe^{II} (0.58 mM), fumitremorgin B (0.58 mM) and α-KG (12 mM). The decay of the Fe–α-KG complex charge transfer band centred at ~520 nm (dashed arrow) and the formation and decay of the spectral feature centred at ~420 nm (arrow) are highlighted. Inset: time-dependent absorbance change at 420 nm. The absorbance reported in **b** was obtained by blanking the spectrometer with the anaerobic buffer containing 100 mM Tris-HCl (pH 7.5). The absorbance reported in the inset was obtained by subtracting the absorbance at 420 nm of the 2 ms spectrum from all other spectra recorded. The trace is the average of two trials. **c**, Spectroscopic evidence for transient radical species. X-band EPR spectra measured at 19 K of reaction samples freeze-quenched at the indicated time points. Measurement conditions: microwave frequency, 9.64 GHz; microwave power, 2 μW; modulation amplitude, 1 mT; and modulation frequency, 100 kHz.

and Extended Data Fig. 9c). The formation and decay of this radical signal closely followed the kinetics of the 420 nm absorption feature observed in stopped-flow optical absorption spectroscopy experiments (Extended Data Fig. 8b), indicating that they are from the same intermediate species. Spin quantification of the EPR signals at ~0.2 s revealed that the Fe^{3+} and radical species accumulated to ~0.35 and ~0.25 equivalents, respectively. The width of the radical EPR signal (~12 mT edge-to-edge width, Fig. 5c) was significantly broader than that of magnetically isolated organic or protein radical signals[29]. Such broadening could be due to a magnetic dipolar interaction of the radical species with an adjacent spin centre, most likely the Fe^{3+} centre depicted in species **D**, **E** or **F** in Fig. 4.

In summary, our FtmOx1 structural and biochemical characterization provides a notable example of the catalytic versatility of mononuclear non-haem iron enzymes and how changes in the secondary coordination sphere to the non-haem iron facilitate unprecedented chemical transformations.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 15 May; accepted 24 August 2015.
Published online 2 November 2015.

1. Casteel, D. A. Peroxy natural products. *Nat. Prod. Rep.* **9,** 289–312 (1992).
2. Casteel, D. A. Peroxy natural products. *Nat. Prod. Rep.* **16,** 55–73 (1999).
3. Chaturvedi, D., Goswami, A., Saikia, P. P., Barua, N. C. & Rao, P. G. Artemisinin and its derivatives: a novel class of anti-malarial and anti-cancer agents. *Chem. Soc. Rev.* **39,** 435–454 (2010).
4. Paddon, C. J. & Keasling, J. D. Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. *Nature Rev. Microbiol.* **12,** 355–367 (2014).
5. Dembitsky, V. M. Bioactive peroxides as potential therapeutic agents. *Eur. J. Med. Chem.* **43,** 223–251 (2008).
6. Widboom, P. F., Fielding, E. N., Liu, Y. & Bruner, S. D. Structural basis for cofactor-independent dioxygenation in vancomycin biosynthesis. *Nature* **447,** 342–345 (2007).
7. Steiner, R. A., Janssen, H. J., Roversi, P., Oakley, A. J. & Fetzner, S. Structural basis for cofactor-independent dioxygenation of *N*-heteroaromatic compounds at the α/β-hydrolase fold. *Proc. Natl Acad. Sci. USA* **107,** 657–662 (2010).
8. Thierbach, S. *et al.* Substrate-assisted O_2 activation in a cofactor-independent dioxygenase. *Chem. Biol.* **21,** 217–225 (2014).
9. Marnett, L. J. Cyclooxygenase mechanisms. *Curr. Opin. Chem. Biol.* **4,** 545–552 (2000).
10. Grundmann, A. & Li, S. M. Overproduction, purification and characterization of FtmPT1, a brevianamide F prenyltransferase from *Aspergillus fumigatus*. *Microbiology* **151,** 2199–2207 (2005).
11. Steffan, N., Grundmann, A., Afiyatullov, S., Ruan, H. & Li, S. M. FtmOx1, a non-heme Fe(II) and α-ketoglutarate-dependent dioxygenase, catalyses the endoperoxide formation of verruculogen in *Aspergillus fumigatus*. *Org. Biomol. Chem.* **7,** 4082–4087 (2009).
12. Kato, N. *et al.* Gene disruption and biochemical characterization of verruculogen synthase of *Aspergillus fumigatus*. *ChemBioChem* **12,** 711–714 (2011).
13. Clifton, I.J. *et al.* Structural studies on 2-oxoglutarate oxygenases and related double-stranded α-helix fold proteins. *J. Inorg. Biochem.* **100,** 644–669 (2006).
14. Hausinger, R. P. Fe(II)/α-ketoglutarate-dependent hydroxylases and related enzymes. *Crit. Rev. Biochem. Mol. Biol.* **39,** 21–68 (2004).
15. Costas, M., Mehn, M. P., Jensen, M. P. & Que, L. Dioxygen activation at mononuclear nonheme iron active sites: Enzymes, models, and intermediates. *Chem. Rev.* **104,** 939–986 (2004).
16. Solomon, E. I. *et al.* Geometric and electronic structure/function correlations in non-heme iron enzymes. *Chem. Rev.* **100,** 235–350 (2000).
17. Kovaleva, E. G. & Lipscomb, J. D. Versatility of biological non-heme Fe(II) centers in oxygen activation reactions. *Nature Chem. Biol.* **4,** 186–193 (2008).
18. Ryle, M. J., Padmakumar, R. & Hausinger, R. P. Stopped-flow kinetic analysis of *Escherichia coli* taurine/α-ketoglutarate dioxygenase: interactions with α-ketoglutarate, taurine, and oxygen. *Biochemistry* **38,** 15278–15286 (1999).
19. Liu, A., Ho, R. Y. N. & Que, L. Alternative reactivity of an α-ketoglutarate-dependent Iron(II) oxygenase: enzyme self-hydroxylation. *J. Am. Chem. Soc.* **123,** 5126–5127 (2001).
20. Volkamer, A., Kuhn, D., Grombacher, T., Rippmann, F. & Rarey, M. Combining global and local measures for structure-based druggability predictions. *J. Chem. Inf. Model.* **52,** 360–372 (2012).
21. Elkins, J. M. *et al.* X-ray crystal structure of *Escherichia coli* taurine/α-ketoglutarate dioxygenase complexed to ferrous iron and substrates. *Biochemistry* **41,** 5185–5192 (2002).
22. Liu, P. *et al.* Protein purification and function assignment of the epoxidase catalyzing the formation of fosfomycin. *J. Am. Chem. Soc.* **123,** 4619–4620 (2001).
23. Vaillancourt, F. H., Yeh, E., Vosburg, D. A., O'Connor, S. E. & Walsh, C. T. Cryptic chlorination by a non-haem iron enzyme during cyclopropyl amino acid biosynthesis. *Nature* **436,** 1191–1194 (2005).
24. Blasiak, L. C., Vaillancourt, F. H., Walsh, C. T. & Drennan, C. L. Crystal structure of the non-haem iron halogenase SyrB2 in syringomycin biosynthesis. *Nature* **440,** 368–371 (2006).
25. Clifton, I. J. *et al.* Crystal structure of carbapenem synthase (CarC). *J. Biol. Chem.* **278,** 20843–20850 (2003).
26. Chang, W. C. *et al.* Mechanism of the C5 stereoinversion reaction in the biosynthesis of carbapenem antibiotics. *Science* **343,** 1140–1144 (2014).
27. Blodgett, J. A. V. *et al.* Unusual transformations in the biosynthesis of the antibiotic phosphinothricin tripeptide. *Nature Chem. Biol.* **3,** 480–485 (2007).
28. Krebs, C., Galonic Fujimori, D., Walsh, C. T. & Bollinger, J. M., Jr. Non-heme Fe(IV)-oxo intermediates. *Acc. Chem. Res.* **40,** 484–492 (2007).
29. Stubbe, J. & van der Donk, W. A. Protein radicals in enzyme catalysis. *Chem. Rev.* **98,** 705–762 (1998).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank H.-w. Liu, S. Elliott and A. Liu for comments on the manuscript. We also thank R. Fan and J. Lee for assistance with the pre-steady state kinetics studies, J. Caradonna for use of stopped-flow instruments, and A. Monzingo for assistance with crystallography software. This work is supported in part by grants from the National Institutes of Health (R01 GM093903 to P.L.; P41 GM104603 to C.E.C.; R01 GM104896 to Y.J.Z.; and R01 GM077387 to M.P.H.), the National Science Foundation (CHE-1309148 to P.L.; CHE-1126268 for the EPR spectrometer), the Welch Foundation (F-1778 to Y.J.Z.), the 973 program (2013CB734000 to L.Z), and Y.G. acknowledges financial support from Carnegie Mellon University. Crystallographic data collection was conducted at advanced light sources (Beamline 5.0.3) and advanced photon sources (BL23-ID-B), Department of Energy (DOE) National User Facility. L.Z. is an awardee of the National Distinguished Young Scholar Program in China (31125002).

Author Contributions P.L., Y.J.Z. and L.Z. designed the study. W.Y. conducted the crystallization experiments and structure determination. H.S., F.S., A.S.H., S.W. and N. N. conducted the biochemical studies. C.-H.W., Y.P. and C.E.C. performed the MS–MS analyses. H.S., Y.G., M.P.H. and A.W. conducted the pre-steady state kinetics and EPR characterization. The manuscript was written by P.L., Y.J.Z. and L.Z. with input from all contributing authors.

Author Information The structural factors and coordinates of FtmOx1 and its complexes with either α-KG or fumitremorgin B have been deposited in the Protein Data Bank with accession codes 4Y5T, 4Y5S and 4ZON. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.L. (pinghua@bu.edu) or Y.J.Z. (jzhang@cm.utexas.edu) or L.Z. (lzhang03@gmail.com).

METHODS

No statistical methods were used to predetermine sample size.

Materials and experimental procedures. Fumitremorgin B was isolated from *Aspergillus fumigatus* strain IM-MF330 according to the procedure summarized in a later section. All reagents were purchased from Sigma-Aldrich unless otherwise stated.

Nuclear magnetic resonance (NMR) spectra were obtained on a Bruker Avance DRX600 spectrometer in the solvents indicated and referenced to residual ^{1}H and ^{13}C signals in deuterated solvents. High-resolution electrospray ionization (ESI) mass spectrometry (MS) measurements were obtained on a Bruker micrOTOF mass spectrometer. High-performance liquid chromatography (HPLC) was performed using an Agilent 1200 Series separations module equipped with Agilent 1200 Series diode array detectors and an Agilent 1200 Series fraction collector, controlled using ChemStation. UV–vis analysis was performed on a Varian Cary 100 Bio UV–vis spectrophotometer.

Sub-cloning and overexpression of wild-type, Y224F- and Y224A-substituted FtmOx1. The coding sequence of the *FtmOx1* gene from *A. fumigatus* Af293 (accession number: XM_742088) was sub-cloned into the EcoRI and XhoI restriction sites of the pASK-IBA3^{+} vector, which places it under the control of the *tet*-promoter and allows for the production of C-terminally strep-tagged FtmOx1. The final recombinant FtmOx1 includes some extra amino acid residues at the N terminus and a strep-tag at the C terminus for purification. The residue numbering used in this manuscript is based on the FtmOx1 sequence deposited in GenBank (accession number: XM_742088). Y224F- and Y224A-substituted FtmOx1 were generated using a Stratagene QuikChange II kit according to the manufacturer's instructions.

Plasmids encoding wild-type, Y224F-, and Y224A-substituted FtmOx1 mutant genes were used to transform *Escherichia coli* BL21 (DE3) cells (Invitrogen Inc.) for protein overexpression. A single colony was used to inoculate a starter culture, which was incubated at 37 °C overnight. Production cultures were grown at 37 °C in Luria–Bertani medium supplemented with 100 μg ml^{-1} ampicillin to an optical density (OD_{600}) of ~0.8 and then cooled to 25 °C. The FtmOx1 protein production was induced by the addition of anhydrotetracycline to a final concentration of 250 μg l^{-1}. The cultures were grown at 25 °C for an additional 16 h before harvesting.

Purification was performed at 4 °C. In a typical purification, ~30 g wet cell paste was resuspended in 100 ml of anaerobic buffer (100 mM Tris-HCl, 50 mM NaCl, and 5 mM 1,10-phenanthroline (pH 7.5)) in an anaerobic Coy chamber. Lysozyme (1.0 mg ml^{-1} final concentration) and DNase I (100 U per gram of cell) were then added into the cell suspension, and the mixture was incubated on ice for 40 min with gentle agitation. The cells were disrupted by sonication (20 cycles of 10 s bursts) using a Fisher Scientific Model 505 Sonic Dismembrator. The supernatant and the cell debris were anaerobically separated by centrifugation at 4 °C for 30 min at 20,000*g*. Streptomycin sulfate was added into the supernatant (~100 ml) to a final concentration of 1% (w/v), and the mixture was incubated on ice for 30 min with gentle agitation. The DNA precipitate was then removed by centrifugation at 20,000*g* for 40 min at 4 °C. The resulting supernatant was mixed with Strep-Tactin resin (50 ml) and incubated on ice for 30 min. After the cell lysate was drained by gravity, the column was washed with washing buffer (100 mM Tris-HCl and 150 mM NaCl (pH 7.5)) until the OD_{260} was <0.05. The FtmOx1 protein was reconstituted by incubating the protein-loaded resin with 50 ml of a solution containing 3.0 mM ammonium ferrous sulfate and 5.0 mM ascorbate at 4 °C for 10 min. After the excess solution was drained by gravity, the resin was further washed with washing buffer until the OD_{280} was <0.05. Recombinant FtmOx1 was eluted with elution buffer (2.5 mM desthiobiotin in 100 mM Tris-HCl and 50 mM NaCl (pH 7.5)). The eluted protein was concentrated, flash frozen with liquid nitrogen, and stored at −80 °C. From 30 g of wet cell paste, ~400 mg of protein was obtained. The purity of the protein was shown by SDS–PAGE (12%) as a single band. The FtmOx1 concentration was calculated using ε_{280} nm of 43,288 M^{-1} cm^{-1} determined by amino acid analysis.

Selenomethionine-incorporated FtmOx1 was prepared using a modified medium. A single colony was used to inoculate 50 ml Luria–Bertani medium supplemented with 100 μg ml^{-1} ampicillin, which was incubated at 37 °C until the OD_{600} was ~0.5. The pre-culture (2 ml) was transferred into 150 ml of minimal media (1 l minimal media contained 50 ml glycerol, 12.8 g Na_2HPO_4–$7H_2O$, 3 g KH_2PO_4, 0.5 g NaCl, 1 g NH_4Cl, 0.2% glucose, 0.1 mM $CaCl_2$, and 2.0 mM $MgSO_4$) supplemented with 100 μg ml^{-1} ampicillin, which was incubated at 37 °C for an additional 5 h. Then, 10 ml of the pre-culture was transferred into 1 l minimal media supplemented with 100 μg ml^{-1} ampicillin and incubated at 37 °C for 12 h. Subsequently, 10 ml 100× amino acid solution mix (100× amino acid solution contained 100 mg lysine, 100 mg threonine, 100 mg phenylalanine, 50 mg leucine, 50 mg isoleucine, and 50 mg valine in 10 ml H_2O) and 100× selenomethionine solution (60 mg L-selenomethionine in 10 ml H_2O) were added to the culture medium. After 0.5 h, the temperature was decreased to 25 °C, and FtmOx1 overexpression was induced by the addition of anhydrotetracycline to a final concentration of 250 μg l^{-1}. The cultures were grown at 18 °C for an additional 12 h before harvesting.

Selenomethionine-incorporated FtmOx1 was purified according to the same procedure described earlier. From 5 g of wet cell paste, ~15 mg of selenomethionine-incorporated FtmOx1 was obtained.

Before crystallization, FtmOx1 was further purified by gel filtration (Superdex 200, GE Healthcare) in buffer containing 100 mM Tris-HCl at pH 7.5 and 50 mM NaCl. After gel filtration, FtmOx1 was concentrated to ~10 mg ml^{-1} and stored at −80 °C for future crystallization experiments.

Isolation of fumitremorgin B. *Aspergillus fumigatus* strain IM-MF330 was isolated from a mud sample collected from the Yellow Sea. A small number of spores growing on a potato dextrose agar slant was inoculated into a 250-ml conical flask containing 40 ml of liquid medium (20% potato infusion, 2.0% glucose, 3.5% sea salt, and distilled water) and then cultured at 28 °C for 3 days on a rotary shaker at 160 rpm. The seed culture (5 ml) was inoculated into 1,000-ml conical flasks, each containing 130 g rice and 80 ml artificial seawater, and incubated without aeration for 19 days. The fermentation product was exhaustively extracted with EtOAc:MeOH (80:20) to yield a crude extract. The crude extract was partitioned between EtOAc and H_2O. The EtOAc layer (10.4 g) was applied to a column of silica gel using a gradient solvent system of 50–100% petroleum ether/CH_2Cl_2 and 0–100% MeOH/CH_2Cl_2 to afford 15 fractions. Fraction MF330F was passed through a Sephadex LH-20 column and eluted with petroleum-ether: CH_2Cl_2:MeOH (5:5:1) to yield 5 sub-fractions. The third fraction MF30F3 was subsequently subjected to HPLC fractionation (Agilent Zorbax SB-C18 5 μm 250 × 9.4 mm column, 3.0 ml min^{-1}, 65% MeOH) to yield verruculogen and fumitremorgin B, respectively.

Crystallization and data collection. FtmOx1 crystallization was set up using the sitting-drop vapour diffusion method by mixing protein and crystallization buffer (100 mM MES (pH 6.5), 50 mM $CoCl_2$, and 2 M ammonium sulfate) at a ratio of 2:1 at room temperature. Sheet-like crystals were visible after 7 days. The FtmOx1 and α-ketoglutarate (α-KG) complex was obtained using both soaking and co-crystallization methods, which led to identical models. Crystal soaking was conducted by transferring the pre-formed FtmOx1 crystals into crystallization mother liquor containing 1 mM α-KG and incubated for 2 h at room temperature. Co-crystallization trials included pre-mixture of the protein with α-KG at a ratio of 1:100 for 2 h before crystallization setup. The crystals were cryoprotected by the addition of 25% glycerol in mother liquor before being vitrified in liquid nitrogen for data collection. To obtain the structure of the FtmOx1–fumitremorigin-B complex, we crystallized FtmOx1 in an anaerobic chamber using identical conditions. Sheet-like crystals appeared within 3 days and continued to grow for another week before reaching maximal size. Fumitremorgin B was dissolved in buffer containing degased crystallization mother liquor with 0.05% TritonX-100 and 20% glycerol to saturation. After centrifugation, to discard insoluble material, the mother liquor containing a saturating amount of fumitremorgin B was used to soak FtmOx1 crystals as sitting drops for 90 min until cryoprotected with degassed mother liquor with 30% glycerol.

Crystal diffraction data were collected at the Advanced Photon Source beamline BL23-ID-B (Argonne, Illinois) for FtmOx1 (wavelength 0.97931 Å) and selenomethionine-incorporated FtmOx1 (wavelength 0.97958 Å). The diffraction data for the FtmOx1–α-KG and FtmOx1–fumitremorgin-B binary complexes were collected at the Advanced Light Source beamline BL5.0.3 (Berkeley, California) at wavelength 0.97648 Å. All data collection were conducted within liquid nitrogen stream at 100 K. The data were processed using the program HKL2000[30]. The statistics for data collection are summarized in Extended Data Table 1.

Structure determination and refinement. The FtmOx1 structure was determined by the single anomalous dispersion method using the selenomethionine data set with phase information to 3.5 Å resolution. The positions of the selenium were determined and refined by Phenix.Autosol[31–33] followed by the density modification program DM in CCP4 suite[34–45]. An initial model was built based on the phase information using the Buccaneer program[40,46,47], further extended and corrected manually by the COOT program[48]. The resolution was extended to the high-resolution limit of 1.95 Å using the native protein data set. Iterative cycles of optimization were performed to improve the quality of the model using the refinement program PHENIX.Refine[49–55], followed by manual rebuilding in COOT[48]. A portion of the of diffraction data (5%) was reserved as an unbiased test set for cross validation (R_{free}) for the model that eventually had an R_{work} of 16.1% and an R_{free} of 19.9%. The structure of the FtmOx1–α-KG binary complex and FtmOx1–fumitremorgin-B complex were both solved by molecular replacement with the FtmOx1 structure as the initial model using Phaser in the CCP4 package[40,56–59]. The co-substrate α-KG and substrate fumitremorgin B were built using COOT followed by several rounds of refinement by PHENIX.Refine[49–55]. For the

FtmOx1–α-KG complex, the final R_{work} was 20.3% with an R_{free} of 26.0%. For the FtmOx1–fumitremorgin-B complex, the final R_{work} was refined to 16.7% and R_{free} to 20.3%. Model quality for all of the structures was evaluated with Ramachandran and MolProbity[51]. The structures show no outlier and most residues were in the favoured region of Ramachandran statistics (98.6% for apo FtmOx1, 98.1% for the FtmOx1–α-KG complex and 98.3% for FtmOx1–fumitremorgin-B). When evaluated by MolProbity, all three structures rank 100% within the specific resolution range (apo FtmOx1 with a MolProbity score of 1.04, FtmOx1–α-KG complex 1.26 and FtmOx1–fumitremorgin-B 1.12, respectively). Refinement statistics are summarized in Extended Data Table 1. Figure 2 and Extended Data Fig. 5 were prepared with PyMol[60].

Oxygen concentration determination for oxygenated buffer. Oxygen-saturated buffer[61] (10 ml) was transferred into syringes (12 cm^3) with a long needle, and the syringe was then sealed. An alkaline KI solution (2.1 M KI and 8.7 M KOH prepared using oxygen-free water) and $MnSO_4$ solution (2.1 M) in oxygen-free water were prepared in a Coy chamber and transferred out using syringes sealed by a rubber septum. The alkaline KI solution (0.2 ml) and the $MnSO_4$ solution (0.2 ml) were quickly aspirated into the syringe containing the 10 ml oxygen saturated buffer. Then, the syringe was quickly sealed again. The sample in the syringe was intensely mixed (turning the syringe ~10 times upside-down until the entire syringe was filled with the floating $Mn(OH)_3$ precipitate); the $Mn(OH)_3$ precipitate formed completely in 45 min according to the following reaction:

$$4Mn^{2+} + O_2 + 8OH^- + 2H_2O \rightarrow 4Mn(OH)_3\downarrow \quad (1)$$

After 45 min, H_2SO_4 solution (0.2 ml, 2.7 M) was aspirated into the syringe, and Mn^{3+} ions oxidized iodide to iodine under acidic conditions according to the following reactions:

$$2Mn(OH)_3(s) + 3H_2SO_4 \rightarrow 2Mn^{3+} + 3SO_4^{2-} + 3H_2O \quad (2)$$

$$2Mn^{3+} + 2I^- \rightarrow 2Mn^{2+} + I_2 \quad (3)$$

Iodine eventually formed I_3^- ions with the excess KI:

$$I_2 + I^- \rightarrow I_3^- \quad (4)$$

The resulting iodine solution was transferred to a sample bottle and immediately titrated with standardized 2.5 mM $Na_2S_2O_3$ solution:

$$I_3^- + 2S_2O_3^{2-} \rightarrow 3I^- + S_4O_6^{2-} \quad (5)$$

According to reactions (1)–(5), one equivalent of oxygen molecule corresponds to four equivalents of $Na_2S_2O_3$. Therefore, the oxygen concentration in the oxygen-saturated buffer was determined based on the amount of standardized 2.5 mM $Na_2S_2O_3$ solution used for titration.

The $Na_2S_2O_3$ concentration was standardized with an iodine solution, which was prepared by mixing a standard KIO_3 solution and KI solution under acidic conditions

$$(KIO_3 + 5KI + 6H^+ \rightarrow 3I_2 + 6K^+ + 3H_2O;\ I_2 + I^- \rightarrow I_3^-;$$
$$I_3^- + 2S_2O_3^{2-} \rightarrow 3I^- + S_4O_6^{2-})$$

α-KG and oxygen stoichiometries in the FtmOx1 catalysis. To examine whether FtmOx1 is capable of catalysing verruculogen oxidation in the absence of other reductants, the FtmOx1 reaction was conducted under the following conditions: a 200 µl anaerobic mixture in 100 mM Tris-HCl (pH 7.5), contained fumitremorgin B (360 µM), α-KG (4 mM), and variable amounts of iron-loaded FtmOx1 (0.25×, 0.5×, 1×, and 2.0× of iron-loaded FtmOx1 relative to the fumitremorgin B concentration). The reaction was initiated by quickly mixing the above solution with 200 µl oxygen-saturated buffer (1.2 mM) in the Coy chamber to make a solution containing 600 µM oxygen. The final reaction mixture contained 180 µM fumitremorgin B, 2 mM α-ketoglutarate, 600 µM oxygen, and a variable amount of iron-loaded FmOx1 (0.25×, 0.5×, 1×, and 2× of FtmOx1 relative to that of verruculogen concentration). After the reaction was initiated, the reaction mixture was sealed and incubated for 0.5 h at 37 °C. The enzymatic reaction was quenched by adding 300 µl chloroform, the precipitated protein was removed by centrifugation at 13,000*g* for 10 min, and the chloroform layer was carefully removed. The reaction mixture was extracted one more time using a second 300-µl volume of chloroform. The combined chloroform layers were concentrated by rotatory evaporation, and the residue was re-dissolved in 100 µl acetonitrile and subjected to HPLC analysis.

To determine α-KG stoichiometry, 200 µl anaerobic reaction mixture (in 100 mM Tris-HCl (pH 7.50)) contained 360 µM fumitremorgin B, 240 µM iron-loaded FtmOx1, and variable amounts of α-KG. The concentration of α-KG was varied to make reaction mixtures containing 0.5×, 1.0×, 1.5×, and 2.0× of α-KG relative to the FtmOx1 concentration. The reaction was initiated by quickly mixing 200 µl of oxygen-saturated buffer (1.2 mM) in the Coy chamber to make a final oxygen concentration of 600 µM. The resulting reaction mixtures contained a final concentration of 180 µM fumitremorgin B, 120 µM iron-loaded FtmOx1, 600 µM oxygen, and variable amounts of α-KG. After initiation, the reaction was sealed and incubated for 0.5 h at 37 °C. The enzymatic reaction was quenched by adding 300 µl chloroform, the precipitated protein was removed by centrifugation at 13,000*g* for 10 min, and the chloroform layer was carefully removed. The reaction mixture was extracted one more time using a second 300-µl volume of chloroform. The combined chloroform layers were concentrated by rotary evaporation, and the residue was re-dissolved in 100 µl of acetonitrile and subjected to HPLC analysis.

To determine oxygen stoichiometry, oxygen-saturated buffer was added to a 600 µl anaerobic reaction mixture (100 mM Tris-HCl (pH 7.5) buffer, 360 µM of fumitremorgin B, 240 µM of iron-loaded FtmOx1, and 480 µM of α-KG). To determine the amount of product formation under 1× of oxygen relative to iron-loaded FtmOx1 concentration, the above reaction mixture was quickly mixed with 120 µl of oxygen-saturated buffer (1.2 mM) in the Coy chamber. To assess the amount of product formation under 2× of oxygen relative to iron-loaded FtmOx1 concentration, the above mixture was quickly mixed with 240 µl of oxygen-saturated buffer (1.2 mM) in the Coy chamber. To determine the amount of production formation under 3× of oxygen relative to iron-loaded FtmOx1 concentration, the above mixture was quickly mixed with 360 µl of oxygen-saturated buffer (1.2 mM) in the Coy chamber. After reaction initiation, the reaction mixtures were sealed and incubated for 0.5 h at 37 °C. The enzymatic reaction was quenched by adding 300 µl chloroform, the precipitated protein was removed by centrifugation at 13,000*g* for 10 min, and the chloroform layer was carefully separated. The reaction mixture was extracted one more time using a second 300-µl volume of chloroform. The combined chloroform layers were concentrated by rotatory evaporation, and the residue was re-dissolved in 100 µl of acetonitrile and subjected to HPLC analysis.

Reactions using Y224F-substituted FtmOx1. For the reactions using Y224F-substituted FtmOx1, the anaerobic reaction mixture (600 µl, in 100 mM Tris-HCl (pH 7.5)) contained 400 µM Y224F-substituted FtmOx1 containing 300 µM Fe^{II}, 400 µM fumitremorgin B, and 1,200 µM α-KG. The reaction was initiated by quickly adding 400 µl of oxygen-saturated buffer (1.2 mM) in the Coy chamber. The resulting reaction mixtures contained a final concentration of 240 µM fumitremorgin B, 240 µM Y224F-substituted FtmOx1 containing 192 µM Fe^{II}, and 720 µM α-KG. After initiation, the reaction mixture was sealed and incubated for 0.5 h at 37 °C. The enzymatic reaction was quenched by adding 300 µl chloroform, the precipitated protein was removed by centrifugation at 13,000*g* for 10 min, and the chloroform layer was carefully separated. The reaction mixture was extracted once more using a second 300-µl volume of chloroform. The combined chloroform layers were concentrated by rotatory evaporation, and the residue was re-dissolved in 100 µl of acetonitrile and subjected to HPLC analysis.

Reactions using Y224A-substituted FtmOx1. Y224A-substituted FtmOx1 was analysed by a procedure similar to that described in 'Reactions using Y224F-substituted FtmOx1', except that Y224A-substituted FtmOx1 was used instead of Y224F-substituted FtmOx1.

HPLC analysis of the FtmOx1 reaction products. Enzymatic reaction products were routinely analysed by HPLC using a Phenomenex reversed phase C18 column (250 mm × 4 mm, 5 µm; Phenomenex). A linear gradient of 30–100% (v/v) acetonitrile in water was run for 30 min with a flow rate of 0.7 ml min^{-1}, followed by 100% (v/v) acetonitrile for 5 min. Before the next injection, the column was equilibrated with 30% (v/v) acetonitrile for 2 min. The separation profile was monitored using a Photo Diode Array detector at 300 nm.

Products of the reaction (compounds **2** and **3**) using wild-type FtmOx1 were characterized by NMR and high-resolution mass spectrometry (see Supplementary Information).

Isolating products from reactions using wild-type, Y224F-, or Y224A-substituted FtmOx1. To characterize the reaction products, a large-scale reaction (270 ml) was performed using purified FtmOx1 (41.1 µM) containing 32 µM Fe^{II}, fumitremorgin B (30.9 µM), and 60 µM α-KG in Tris-HCl buffer (50 mm, pH 7.0) at 30 °C for 2 h. Chloroform (300 ml) was added to the reaction mixture. The chloroform layer was transferred into centrifuge bottles and centrifuged at 5,000 rpm for 10 min to remove precipitated proteins. The reaction mixture was extracted once more using a second 300-ml volume of chloroform. The combined chloroform layers were dried over Na_2SO_4 for 0.5 h and concentrated by rotary evaporation. The residue was subjected to HPLC separation on a C18 column (4.6 × 150 mm). A linear gradient of 30–100% (v/v) acetonitrile in water was run for 25 min with a flow rate of 0.7 ml min^{-1}, followed by 100% (v/v) acetonitrile for 5 min. Before the next injection, the column was equilibrated with 30% (v/v)

acetonitrile for 2 min. The elution was monitored using Photo Diode Array detector at 300 nm.

Products of the reaction using Y224F- and Y224A-substituted FtmOx1 variants (compounds **2**–**6**) were characterized using NMR and high-resolution mass spectrometry (see Supplementary Information).

Determining the α-KG dissociation constant. To maintain an anaerobic environment, all of the solutions were made anaerobic by several rounds of freeze-pump-thaw degassing. All spectroscopic studies used a 1-cm light path cuvette. After blanking against FtmOx1, spectra were recorded for samples to which anaerobic α-KG had been added. Titration plots were obtained by plotting the absorption at 520 nm. The titration data were fitted to equation (6)[62]

$$A_{obs} = A_{max}[\text{E–L}] \,/\, n[\text{FtmOx1}_T] \quad (6)$$

in which the observed absorption (A_{obs}) was equal to the maximal absorption (A_{max}) multiplied by the concentration of enzyme–ligand complex ([E–L]) divided by the concentration of ligand binding sites (the number (n) of ligands bound per subunit multiplied by the total concentration of FtmOx1 containing Fe^{II} ([$FtmOx1_T$])). The concentration of enzyme–ligand complex was obtained using equation (7)

$$[\text{E}-\text{L}] = \left\{ \begin{matrix} (K_d + [\alpha - \text{KG}_T] + n[\text{FtmOx1}_T]) \\ \pm\sqrt{(K_d + [\alpha - \text{KG}_T] + n[\text{FtmOx1}_T])^2 - 4[\alpha - \text{KG}_T]n[\text{FtmOx1}_T]} \end{matrix} \right\} \Big/ 2 \quad (7)$$

where K_d is the apparent ligand affinity and [α-KG_T] is the total α-KG concentration. The K_d values were determined from equations (6) and (7) using nonlinear curve fitting (OriginPro 8 software).

Self-hydroxylation in wild-type FtmOx1 protein. The FtmOx1 wild-type protein (0.6 mM) was mixed with (2.0 mM) α-KG in the anaerobic Coy chamber to form the pink species, and the UV–vis spectrum was recorded anaerobically using an S.I. Photonics CCD-440 spectrophotometer. All spectroscopic studies used a 1-cm light path cuvette. Upon exposing the above solution to O_2, the solution slowly changed to a blue colour, and the process was monitored using a Cary Bio UV–vis spectrometer.

Self-hydroxylation in FtmOx1(Y224F) variant. The Y224F-substituted FtmOx1 (1.1 mM) was mixed with α-KG (4.0 mM) in the anaerobic Coy chamber to form the binary complex (a pink species), and the UV spectrum was monitored anaerobically using an S.I. Photonics CCD-440 spectrophotometer. All spectroscopic studies used a 1-cm light path cuvette. Upon exposing the above solution to O_2, the solution slowly changed to a blue colour, and the process was monitored using a Cary Bio UV–vis spectrometer.

MS/MS analysis of FtmOx1. The following protein samples were analysed by tandem MS: (1) wild-type FtmOx1; (2) wild-type FtmOx1 treated with α-KG and oxygen; (3) Y224F-substituted FtmOx1; (4) Y224F-substituted FtmOx1 treated with α-KG and oxygen; and (5) Y224F-substituted FtmOx1 after single-turnover experiments in the presence of fumitremorgin B, α-KG, and oxygen. These protein samples (~1.5 nmol) were dissolved in 50 mM ammonium bicarbonate (pH 8.0) buffer to make a 50 µl solution. Trypsin Gold (Promega US) was added to these solutions in a 1:50 (w/w) ratio, and the proteins were digested for 18 h at 37 °C. A C18 Ziptip (Millipore) was then used to desalt each peptide sample. Each digested sample (500 fmol) was injected and analysed by liquid chromatography (LC)–MS/MS on either an LTQ-Orbitrap XL mass spectrometer or a Q Exactive Plus Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific) coupled with a Triversa Nanomate system (Advion Biosystems, Inc.), and a nanoACQUITY UPLC (Waters) with C18 reversed phase trap (2G-V/MTrap 5 µm Symmetry C18 180 µm × 20 mm) and analytical (1.7 µm BEH130 C18 150 µm × 100 mm) columns. Mobile phase A consisted of 98:2 water:ACN with 0.1% formic acid (FA) and mobile phase B contained 98:2 ACN:water with 0.1% FA. Peptide samples were loaded into the trap column at 2% B with a flow rate of 4.0 µl min^{-1} for 4 min and then transferred to the analytical column at 0.5 µl min^{-1}. The gradient was increased to 40% B over 40 min. For tandem MS analyses, data-dependent high resolution higher-energy collisional dissociation mass spectra were acquired in the Orbitrap mass analysers and Xcalibur was used for data analysis. To identify target peptides, the ProteinProspector program from University of California, San Francisco was used to predict the potential product ions and match them to MS/MS product ions. The mass spectra were manually examined to verify the assignments.

Pre-steady-state characterization of FtmOx1. Stopped-flow experiments were performed on an Applied Photophysics SX20 stopped-flow spectrometer operating in an MBraun UNilab glove box. To maintain an anaerobic environment, all of the solutions were prepared in an inert atmosphere box. An oxygen-saturated buffer solution (100 mM Tris-HCl (pH 7.5)) was mixed with an equal volume of an oxygen-free solution containing FtmOx1 (0.65 mM), Fe^{II} (0.58 mM), α-KG (12 mM), substrate (0.58 mM), and 20% glycerol to initiate the reaction. Absorbance scans from 300 to 700 nm were collected with a diode-array detector at 8 °C. The resulting data were processed using SigmaPlot software.

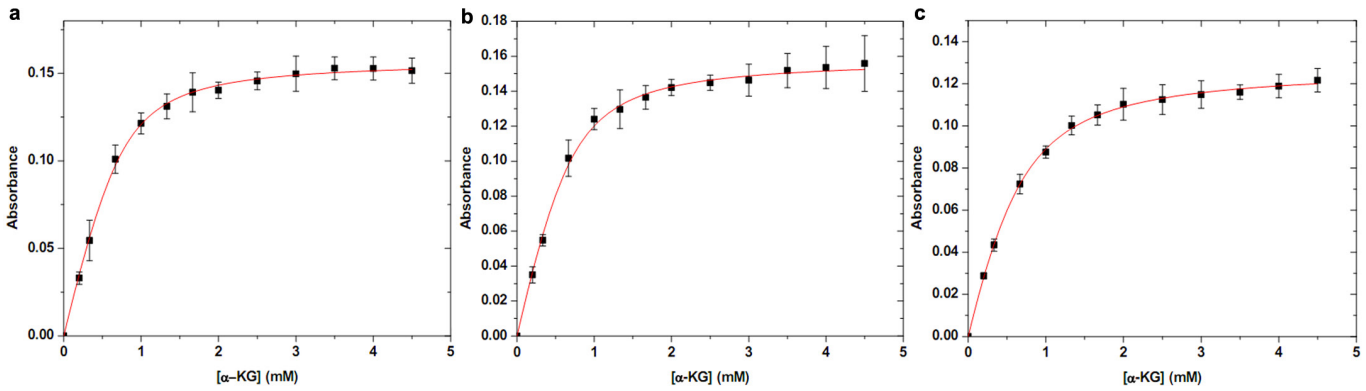
Freeze-quench experiments were performed using a KinTek quench-flow instrument. Analogous to the stopped-flow experiments, an oxygen-saturated buffer solution (100 mM Tris-HCl (pH 7.5)) was mixed with an equal volume of an oxygen-free solution containing FtmOx1 (0.65 mM), Fe^{II} (0.58 mM), α-KG (12 mM), substrate (0.58 mM), and 20% glycerol to initiate the reaction at 8 °C. The resulting reaction was terminated by injection of the solution into liquid ethane (–90 °C) at various time points. The reaction time of a freeze-quenched sample is the sum of the ageing time and the quench time. The ageing time was the transit time for the reaction mixture through the ageing hose. The quench time corresponded to the time required after injection into the cryosolvent for the reaction mixture to be cooled sufficiently to prevent further reaction and was estimated as ~5 ms (ref. 63).

The chemical-quench-flow experiments were performed using a KinTek quench-flow instrument. Analogous to the freeze-quench experiments, an oxygen-saturated buffer solution (100 mM Tris-HCl (pH 7.5)) was mixed with an equal volume of an oxygen-free solution containing FtmOx1 (0.65 mM), Fe^{II} (0.58 mM), α-KG (12 mM), substrate (0.58 mM), and 20% glycerol to initiate the reaction at 8 °C. The resulting reaction was terminated by injecting the solution into a microcentrifuge tube containing 4× volumes of acetone at the desired reaction times. Before HPLC analysis, the samples were centrifuged to remove protein, and the supernatant was concentrated by rotatory evaporation. The concentrated samples were subjected to HPLC separation on a C18 column (4.6 × 100 mm). A linear gradient of 30–100% (v/v) acetonitrile in water was run for 25 min with a flow rate of 0.5 ml min^{-1}, followed by 100% (v/v) acetonitrile for 3 min. Before the next injection, the column was equilibrated with 30% (v/v) acetonitrile for 2 min. The substances were detected with a Photo Diode Array detector at 300 nm.

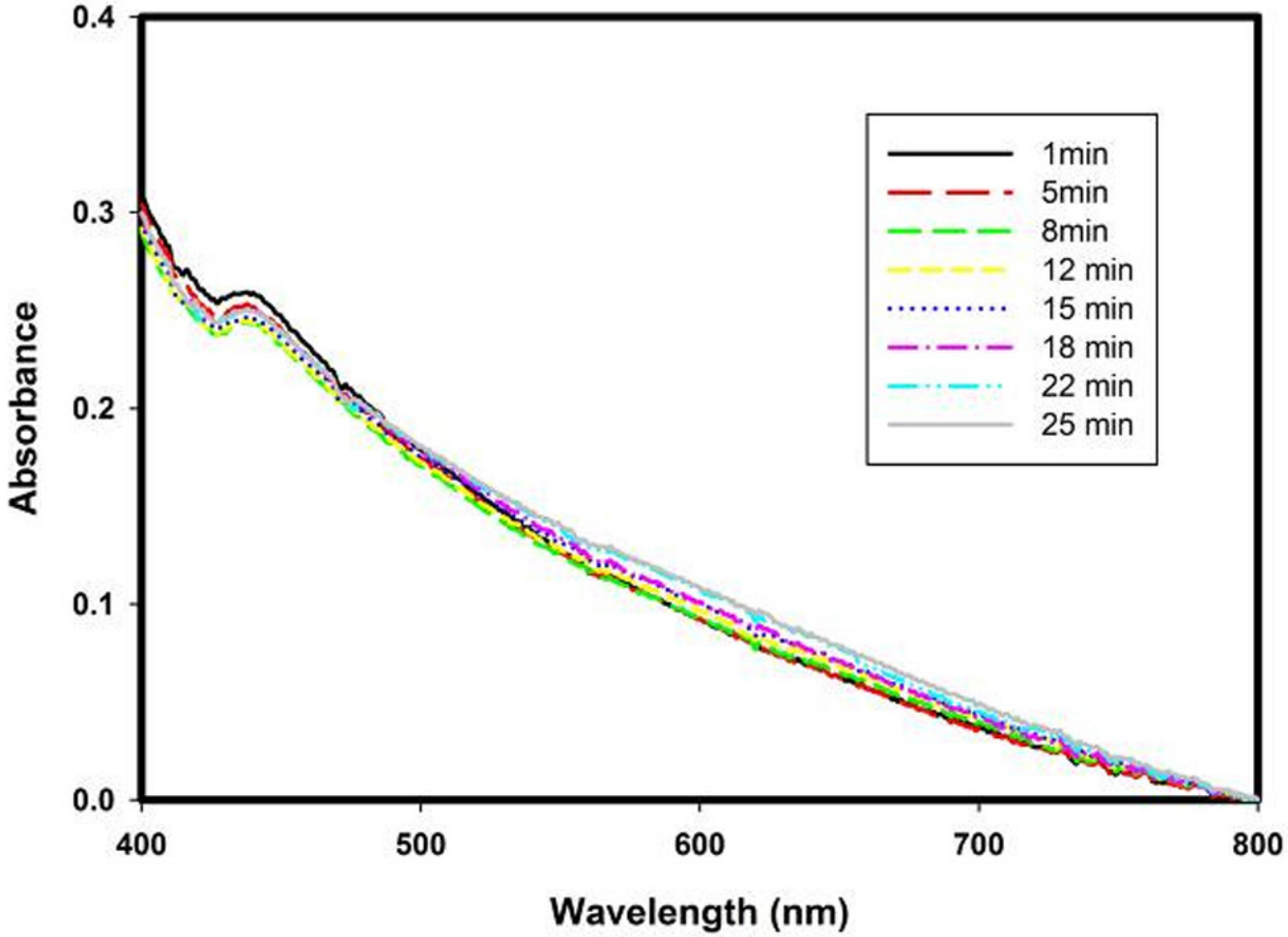
X-band (9.64 GHz) EPR spectra were recorded on a Bruker E500A spectrometer equipped with an Oxford ESR 910 cryostat for low-temperature measurements. The microwave frequency was calibrated with a frequency counter, and the magnetic field was calibrated with an NMR gaussmeter. The temperature of the X-band cryostat was calibrated with a carbon-glass resistor temperature probe (CGR-1-1000 LakeShore Cryotronics). For all EPR spectra, a modulation frequency and amplitude of 100 kHz and 1 mT were used. The EPR spectral simulations were performed using the simulation software Spin Count developed by one of the authors.

30. Minor, W. & Otwinowski, Z. in *Methods in Enzymology, Macromolecular Crystallography* (Academic Press, 1997).
31. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66,** 213–221 (2010).
32. Terwilliger, T. C. Maximum-likelihood density modification. *Acta Crystallogr. D* **56,** 965–972 (2000).
33. Terwilliger, T. C. *et al.* Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr. D* **65,** 582–601 (2009).
34. Baker, D., Bystroff, C., Fletterick, R. J. & Agard, D. A. PRISM: topologically constrained phased refinement for macromolecular crystallography. *Acta Crystallogr. D* **49,** 429–439 (1993).
35. Bricogne, G. Geometric sources of redundancy in intensity data and their use for phase determination. *Acta Crystallogr. A* **30,** 395–405 (1974).
36. Brünger, A. T. Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355,** 472–475 (1992).
37. Cowtan, K. Error estimation and bias correction in phase-improvement calculations. *Acta Crystallogr. D* **55,** 1555–1567 (1999).
38. Cowtan, K. D. & Main, P. Improvement of macromolecular electron-density maps by the simultaneous application of real and reciprocal space constraints. *Acta Crystallogr. D* **49,** 148–157 (1993).
39. Cowtan, K. D. & Main, P. Phase combination and cross validation in iterated density-modification calculations. *Acta Crystallogr. D* **52,** 43–48 (1996).
40. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D* **67,** 235–242 (2011).
41. Sayre, D. Least-squares phase refinement. II. High-resolution phasing of a small protein. *Acta Crystallogr. A* **30,** 180–184 (1974).
42. Schuller, D. J. MAGICSQUASH: more versatile non-crystallographic averaging with mulitple constraints. *Acta Crystallogr. D* **52,** 425–434 (1996).
43. Swanson, S. M. Core tracing: depicting connections between features in electron density. *Acta Crystallogr. D* **50,** 695–708 (1994).
44. Wang, B. C. Resolution of phase ambiguity in macromolecular crystallography. *Methods Enzymol.* **115,** 90–112 (1985).
45. Zhang, K. & Main, P. The use of Sayre's equation with solvent flattening and histogram matching for phase extension and refinement of protein structures. *Acta Crystallogr. A* **46,** 377–381 (1990).

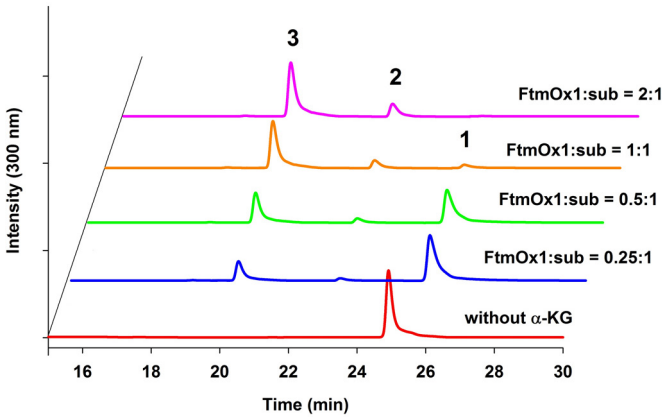
46. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D* **62,** 1002–1011 (2006).
47. Cowtan, K. Fitting molecular fragments into electron density. *Acta Crystallogr. D* **64,** 83–89 (2008).
48. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66,** 486–501 (2010).
49. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D* **68,** 352–367 (2012).
50. Berkholz, D. S., Shapovalov, M. V., Dunbrack, R. L., Jr & Karplus, P. A. Conformation dependence of backbone geometry in proteins. *Structure* **17,** 1316–1325 (2009).
51. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66,** 12–21 (2010).
52. Headd, J. J. *et al.* Use of knowledge-based restraints in phenix.refine to improve macromolecular refinement at low resolution. *Acta Crystallogr. D* **68,** 381–390 (2012).
53. Moriarty, N. W., Grosse-Kunstleve, R. W. & Adams, P. D. electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr. D* **65,** 1074–1080 (2009).
54. Tronrud, D. E., Berkholz, D. S. & Karplus, P. A. Using a conformation-dependent stereochemical library improves crystallographic refinement of proteins. *Acta Crystallogr. D* **66,** 834–842 (2010).
55. Urzhumtseva, L., Afonine, P. V., Adams, P. D. & Urzhumtsev, A. Crystallographic model quality at a glance. *Acta Crystallogr. D* **65,** 297–300 (2009).
56. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674 (2007).
57. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. Likelihood-enhanced fast translation functions. *Acta Crystallogr. D* **61,** 458–464 (2005).
58. Read, R. J. Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Crystallogr. D* **57,** 1373–1382 (2001).
59. Storoni, L. C., McCoy, A. J. & Read, R. J. Likelihood-enhanced fast rotation functions. *Acta Crystallogr. D* **60,** 432–438 (2004).
60. Schrodinger, LLC. *The PyMOL Molecular Graphics System, Version 1.3r1* (2010).
61. Helm, I., Jalukse, L., Vilbaste, M. & Leito, I. Micro-Winkler titration method for dissolved oxygen concentration measurement. *Anal. Chim. Acta* **648,** 167–173 (2009).
62. Ryle, M. J., Padmakumar, R. & Hausinger, R. P. Stopped-flow kinetic analysis of *Escherichia coli* taurine/α-ketoglutarate dioxygenase: interactions with α-ketoglutarate, taurine, and oxygen. *Biochemistry* **38,** 15278–15286 (1999).
63. Baldwin, J. *et al.* Mechanism of rapid electron transfer during oxygen activation in the R2 subunit of *Escherichia coli* ribonucleotide reductase. 1. Evidence for a transient tryptophan radical. *J. Am. Chem. Soc.* **122,** 12195–12206 (2000).



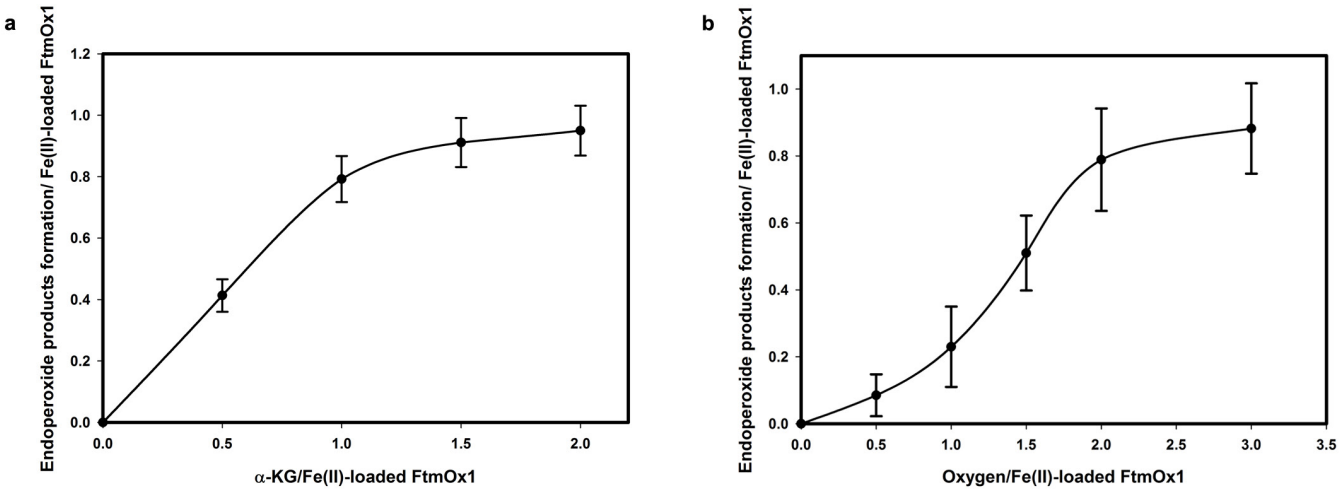
Extended Data Figure 1 | Characterization of FtmOx1–α-KG complex. **a**, Wild-type FtmOx1 and α-KG binding curve. The increase in absorbance at 520 nm as a function of α-KG concentration when it was added to a solution of wild-type FtmOx1 (0.9 mM) and Fe^{II} (0.72 mM) is plotted. On the basis of the equations described in the Methods (determining the α-KG dissociation constant), the K_d for wild-type FtmOx1 and α-KG is ~185 ± 35 μM. **b**, Y224F-substituted FtmOx1 and α-KG binding curve. The increase of absorbance at 520 nm as a function of α-KG concentration when it was added to a solution of Y224F-substituted FtmOx1 (0.9 mM) and Fe^{II} (0.7 mM) is plotted. K_d for Y224F-substituted FtmOx1 and α-KG is ~198 ± 58 μM. **c**, Y224A-substituted FtmOx1 and α-KG binding curve. The increase of absorbance at 520 nm as a function of α-KG concentration when it was added to a solution of Y224A-substituted FtmOx1 (0.7 mM) and Fe^{II} (0.51 mM) is plotted. K_d for Y224A-substituted FtmOx1 and α-KG is 204 ± 43 μM. In **a**–**c**, the K_d was calculated based on the concentration of iron-loaded FtmOx1. The experiments were replicated three times and error bars represent s.e.m.



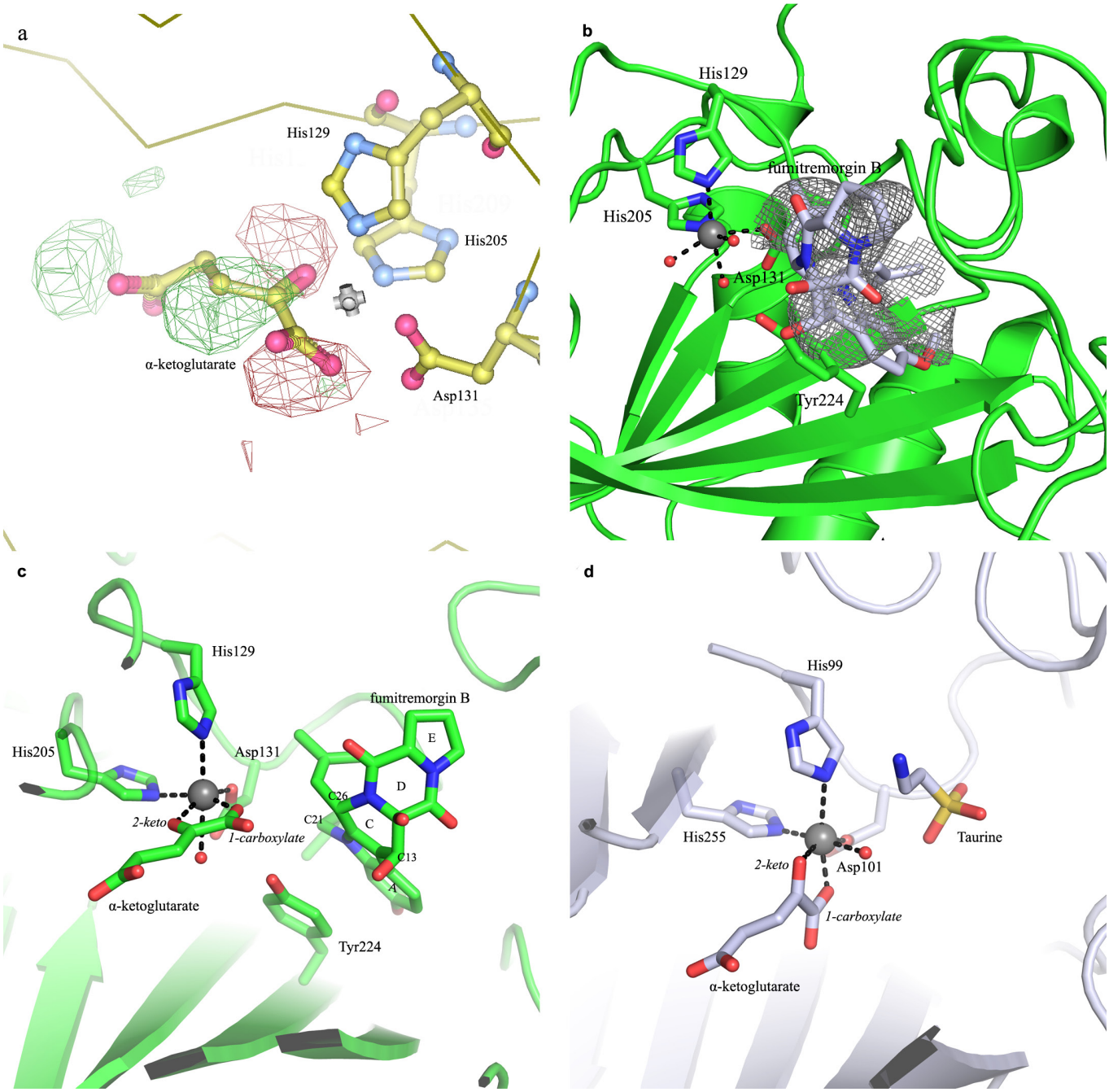
Extended Data Figure 2 | Suppression of DOPA formation by the presence of substrate fumitremorgin B. There is no immediate evidence for the formation of DOPA upon the exposure of the FtmOx1–α-KG complex to O_2 when the substrate fumitremorgin B is present. Spectra were recorded after FtmOx1 was used as the control to blank the UV–visible absorption reading.



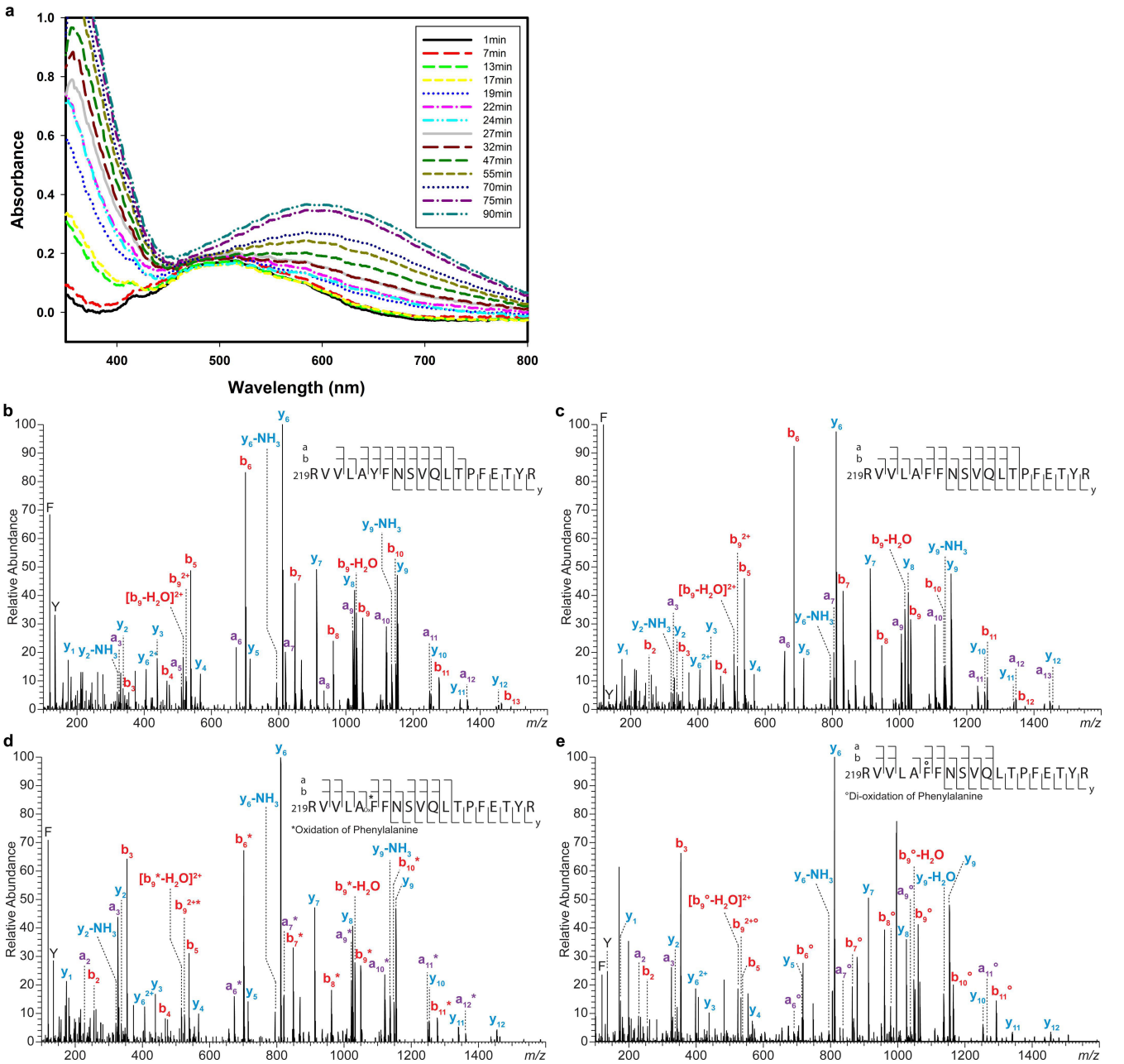
Extended Data Figure 3 | HPLC chromatograms of the FtmOx1 reaction enzyme-concentration dependence. Chromatograms of FtmOx1 reactions with increasing amounts of FtmOx1 relative to the amount of substrate. The reaction mixture contained 100 mM Tris-HCl, (pH 7.5), 180 μM fumitremorgin B, 2 mM α-ketoglutarate, and variable amounts of FmOx1. Identities of the peaks were assigned based on subsequent NMR and MS characterizations of the isolated compounds. This experiment indicates that FtmOx1 is capable of catalysing endoperoxides formation in the absence of any other reductants.



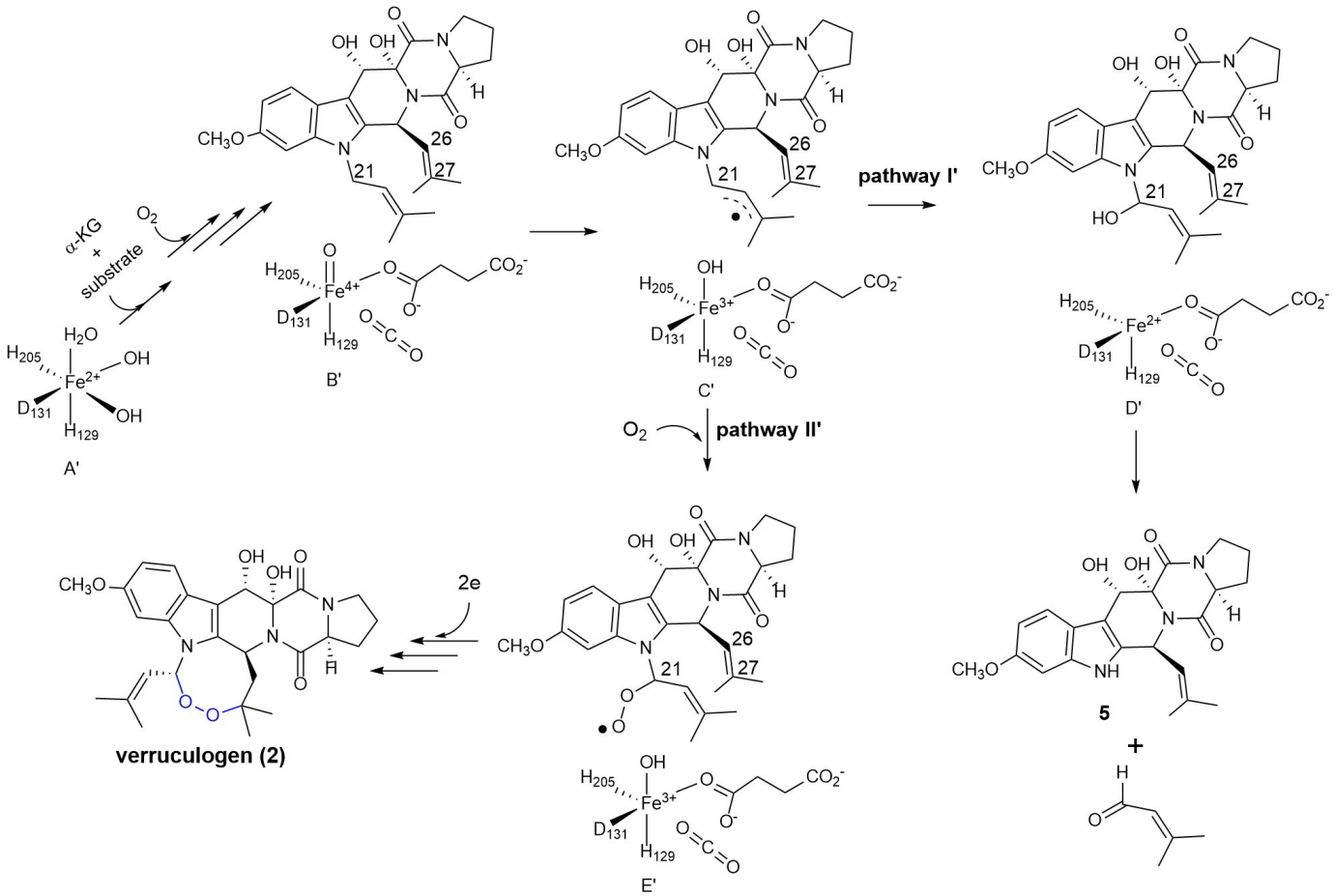
Extended Data Figure 4 | Stoichiometry determination for α-KG and O_2 in FtmOx1 reaction. a, **b**, Equivalents of endoperoxide products (**2** and **3**) produced as a function of the ratio of α-KG to iron-loaded FtmOx1 (**a**) and oxygen to iron-loaded FtmOx1 (**b**). The quantification was conducted based on the fumitremorgin B (**1**), compound **2**, and compound **3** internal standards. All calculations were based on the concentration of iron-loaded FtmOx1. The experiments were replicated three times and error bars represent s.d.



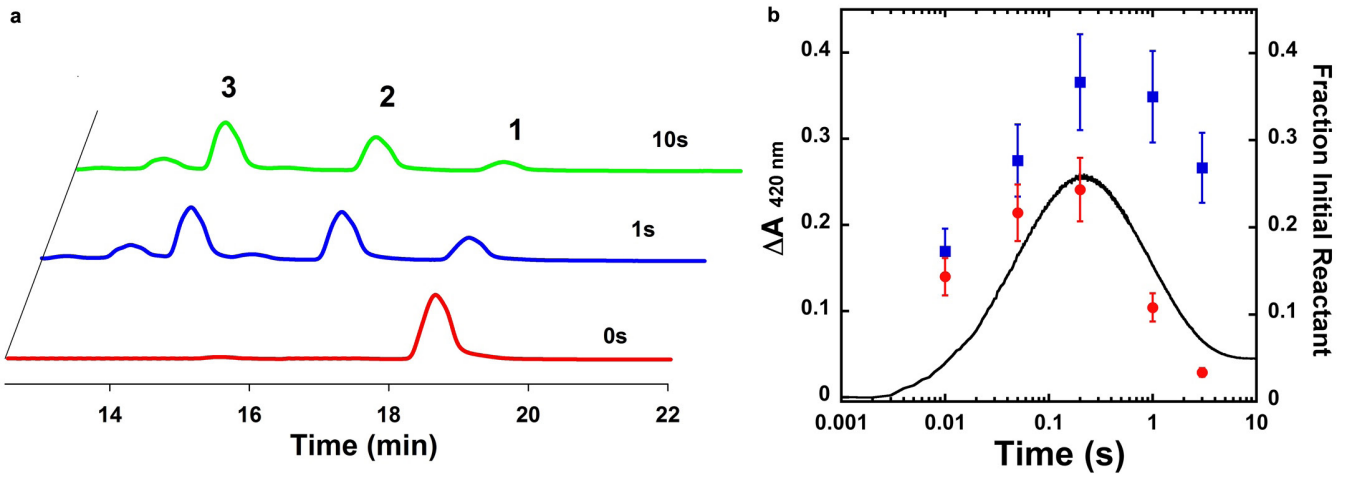
Extended Data Figure 5 | Structural comparison of the active site topologies between FtmOx1 and TauD. a, Examination of the alternative configuration of α-KG in the FtmOx1–α-KG binary complex using the configuration of α-KG in the TauD–α-KG binary complex. We modelled α-KG in this alternative binding mode and calculated the difference map. In the $F_o - F_c$ map, strong positive density (green) and negative density (red) are shown even when contoured to high level (3.3σ), indicating that this configuration is not correct for the FtmOx1–α-KG complex. **b**, The $F_o - F_c$ map at the active site of the FtmOx1–fumitremorgin-B complex. A model of the substrate fumitremorgin B is superimposed onto the difference map, which is contoured at 2.8σ. **c**, Side-by-side comparison of FtmOx1 and TauD active-site topologies. In the left panel, the superimposition of the binary structures of FtmOx1–α-KG and FtmOx1–fumitremorgin-B (**1**) show that the remaining site for oxygen binding and activation is blocked from the substrate by Y224. **d**, In contrast, in the structure of the TauD–taurine–α-KG tertiary complex, the remaining site for O_2 binding and activation directly faces the substrate (taurine).



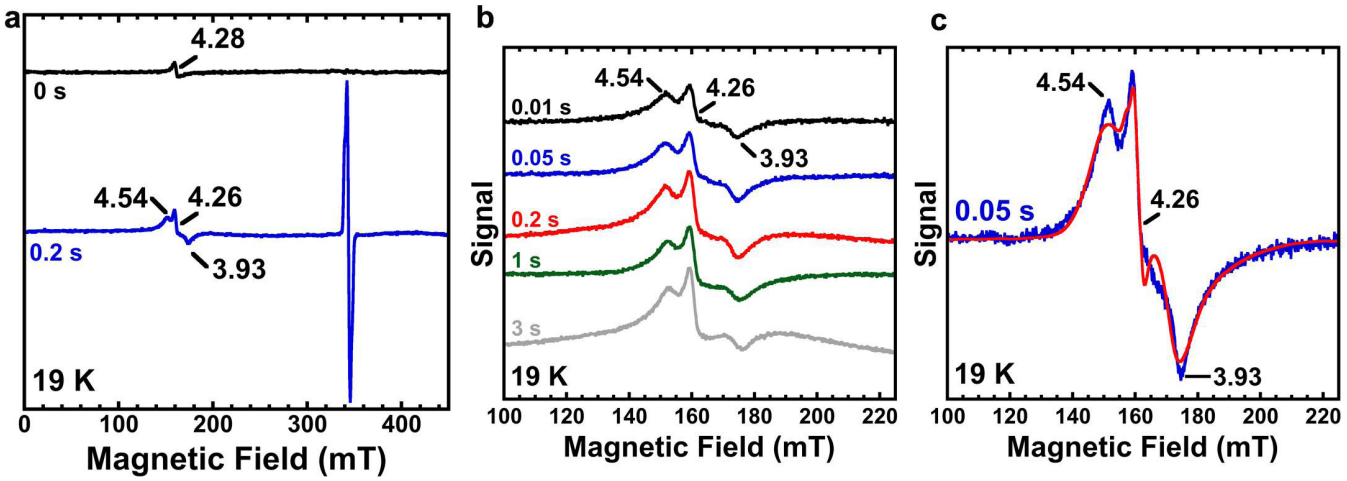
Extended Data Figure 6 | Characterization of FtmOx1 Y224F variant. **a**, Self-hydroxylation reaction in Y224F-substituted FtmOx1. Formation of DOPA upon exposure of the Y224F-substituted FtmOx1–α-KG complex to O_2. **b**–**e**, MS/MS analyses of Y224F-substituted FtmOx1. **b**, MS/MS spectrum of the triply charged parent ion at *m/z* 768.4109 of a tryptic digested peptide (residue 219–237) from wild-type FtmOx1. **c**, MS/MS spectrum of the triply charged parent ion at *m/z* 763.0793 of a tryptic digested peptide (residue 219–237) from Y224F-substituted FtmOx1. **d**, MS/MS spectrum of the triply charged parent ion at *m/z* 768.4109 of a tryptic digested peptide (residue 219–237) after exposure Y224F(FtmOx1)–α-KG tertiary complex to O_2. **e**, MS/MS spectrum of the triply charged parent ion at *m/z* 773.7426 of a tryptic digested peptide (residue 219–237) for DOPA formed upon exposure of FtmOx(Y224F)–α-KG complex to O_2 in the absence substrate fumitremorgin B.



Extended Data Figure 7 | Mechanistic model for the production of dealkylation products in FtmOx1 Y224A or Y224F variants.



Extended Data Figure 8 | Pre-steady-state analyses of FtmOx1 reactions. **a**, HPLC chromatograms for FtmOx1 reactions chemically quenched at the indicated times. The reaction mixture in 100 mM Tris-HCl (pH 7.5) buffer contained FtmOx1 (0.65 mM), Fe^{II} (0.58 mM), α-KG (12 mM), substrate (0.58 mM), and 20% glycerol. The mixture was mixed with O_2-saturated buffer to initiate the reaction. There is an extra signal next to compound **3**, which might be due to other chemicals released during the quench process. Results from the chemical quench experiment indicate that FtmOx1 catalysis is on the timescale of a few seconds per cycle. **b**, Time-dependent 420 nm absorption change (black solid curve) determined by stopped-flow optical absorption spectroscopy and the concentrations of the high-spin Fe^{3+} species (blue squares) and the $g = 2$ species (red dots) determined in the rapid-freeze-quench EPR experiments. The black solid curve is associated with the left y axis and is from the average of two stopped-flow trials. The blue squares and red dots are associated with the right y axis and are from the average of two rapid-freeze-quench EPR experiments. The experiments were repeated twice, and error bars reflect the uncertainty of the packing factor of rapid-freeze-quench EPR samples, which is around ±10%.



Extended Data Figure 9 | EPR spectroscopic analyses of FtmOx1 reactions. a, X-band EPR spectra measured at 19 K in reaction samples prepared at the indicated times. The black line shows the sample containing the FtmOx1–Fe^{II}–α-KG complex in the absence of O_2. (There is a very small signal at $g \approx 4.3$ region, only accounted for by $<5\,\mu M$ iron in the sample, which might be due to a very small amount of Fe^{3+} from inactive enzyme.) Bottom, the reaction sample freeze-quenched at ~0.2 s after mixing the FtmOx1–Fe^{II}–α-KG complex with O_2. It has two signals: an Fe^{3+} ($g = 4.54$, 4.26, and 3.93) and a radical signal at the $g = 2$ region. **b**, X-band EPR spectra measured at 19 K for samples freeze-quenched at the indicated times showing the formation of high-spin ferric species on the time scale within 1 s. The reaction was initiated by mixing the FtmOx1–Fe^{II}–α-KG complex with O_2. *g*-values are indicated in the figure. **c**, X-band EPR spectra measured at 19 K for samples freeze-quenched at 0.05 s and the spectral simulation for an $S = 5/2$ high-spin ferric species. The simulation parameters are: $D = 0.3\ cm^{-1}$, $E/D = 0.266$, $\sigma(E/D) = 0.03$, and $g = 4.54$, 4.26, 3.93. Measurement conditions in **a**–**c**: microwave frequency, 9.64 GHz; microwave power, 0.2 mW; modulation amplitude, 1 mT; and modulation frequency, 100 kHz.

Extended Data Table 1 | X-ray crystallography data collection and refinement statistics

	FtmOx1		FtmOx1•α-KG Complex	FtmOx1•fumitremorgin B complex
Data collection	Set-Met	Native		
Space group	P 1 2_1 1	P 1 2_1 1	P 1 2_1 1	P 1 2_1 1
Cell dimensions				
a, *b*, *c* (Å)	60.6, 45.8, 105.2	60.4, 45.6, 105.4	60.6, 45.8, 105.4	60.3, 45.4, 104.8
α, β, γ (°)	90.0, 100.5, 90.0	90.0, 99.7, 90.0	90.0, 100.0, 90.0	90.0, 100.3, 90.0
Resolution (Å)	48.23 - 3.49 (3.63 - 3.51) *	42.91 - 1.95 (2.02 - 1.95)	36.13 - 2.54 (2.63 - 2.54)	36.05 - 2.11 (2.19 - 2.11)
R_{sym}	0.114 (0.157)	0.101 (0.725)	0.125 (0.691)	0.077 (0.338)
$I/\sigma I$	13.19 (7.64)	18.33 (2.02)	10.53 (1.50)	17.08 (2.80)
Completeness (%)	99.90 (100.00)	99.90 (99.22)	99.93 (99.35)	99.86 (98.63)
Redundancy	6.1 (3.2)	6.6 (5.7)	3.7 (3.7)	3.7 (3.1)
Refinement				
Resolution (Å)		42.91 - 1.95 (2.02 - 1.95)	36.13 - 2.54 (2.63 - 2.54)	36.05 - 2.11 (2.19 - 2.11)
No. reflections		41711	18951	32435
R_{work}/ R_{free}		0.1643/0.2043	0.1756/0.2340	0.1670/0.2033
No. atoms		4906	4704	4884
Protein		4535	4535	4535
Ligand/ion		25	23	38
Water		346	146	311
B-factors (Å^2)		29.5	35.8	32.4
Protein		29.2	35.8	31.9
Ligand/ion		47.5	41.0	49.6
Water		32.0	35.3	37.2
R.m.s deviations				
Bond lengths (Å)		0.005	0.017	0.004
Bond angles (°)		0.97	0.93	0.87

*Highest resolution shell is shown in parenthesis.

CORRIGENDUM

doi:10.1038/nature15533

Corrigendum: A basal ichthyosauriform with a short snout from the Lower Triassic of China

Ryosuke Motani, Da-Yong Jiang, Guan-Bao Chen, Andrea Tintori, Olivier Rieppel, Cheng Ji & Jian-Dong Huang

Nature **517,** 485–488 (2015); doi:10.1038/nature13866

The data matrix in the original Supplementary Data 3 of this Letter reproduced the tree topology shown in Extended Data Fig. 3 but the accompanying character descriptions did not match the coding given in the data matrix. (The numbering of character states was shifted by 1 because of a typo that occurred while editing the list in a spreadsheet, and the character state 3, which was erroneously numbered 4, was accidentally omitted from the list.) The Supplementary Information accompanying this Corrigendum contains the revised Supplementary Data 3. The revised Supplementary Data 3 also reproduces the tree topology shown in Extended Data Fig. 3 of the original Letter; note that the characters have been reordered in the revised Supplementary Data 3 for anatomical consistency.

In addition, the tree statistics originally published in the Extended Data Fig. 3 legend were wrong because they were derived from a matrix where *Parvinatator* was removed from the original matrix. The correct statistics reflecting all 56 taxa in the original matrix are 243 equally most parsimonious trees of TL = 529, CI = 0.423 and RI = 0.796.

Supplementary Information is available in the online version of this Corrigendum.

CORRIGENDUM

doi:10.1038/nature15737

Corrigendum: Influence maximization in complex networks through optimal percolation

Flaviano Morone & Hernán A. Makse

Nature **524,** 65–68 (2015); doi:10.1038/nature14604

In the Acknowledgements section of this Letter, 'ARL' should read 'Army Research Laboratory Cooperative Agreement Number W911NF-09-2-0053 (the ARL Network Science CTA)'. This has been corrected in the online versions of the paper.

CORRECTIONS & AMENDMENTS

RETRACTION

doi:10.1038/nature15745

Retraction: Non-blinking semiconductor nanocrystals

Xiaoyong Wang, Xiaofan Ren, Keith Kahen, Megan A. Hahn, Manju Rajeswaran, Sara Maccagnano-Zacher, John Silcox, George E. Cragg, Alexander L. Efros & Todd D. Krauss

Nature **459,** 686–689 (2009); doi:10.1038/nature08072

In this Letter, we reported the unusual non-blinking characteristics of the fluorescence from individual CdZnSe/ZnSe alloyed quantum dots. However, it has recently come to our attention that similar fluorescence behaviour was seen by Celso de Mello Donega, Daniel Vanmaekelbergh and co-workers from a single fluorophore on bare silica glass. In particular, individual fluorescence spots from single molecules were found to be non-blinking, and fluorescence spectra looked similar to what we reported in our Letter. We corroborated their findings by conducting experiments of our own on bare quartz coverslips, and on quartz coverslips coated with polymethyl methacrylate (PMMA). Although these same control experiments were performed by us before publication, this time we clearly observed non-blinking fluorescence from isolated spots on the coverslip. Furthermore, the fluorescence spectra from these spots were in all practical respects identical to what we reported in our Letter. Subsequent investigations by us have revealed that the surprising origins of the unusual fluorescence come from individual, molecular defects in silica glasses, brightened by the polymer coating. The details of these new findings will be the subject of future publications[1]. After examining the data of de Mello Donega and colleagues, and determining that we were both observing the same phenomena, we concluded that we cannot attribute the fluorescence we observed to CdZnSe/ZnSe quantum dots. In view of these new results, we therefore wish to retract the paper and sincerely apologize for our error. All authors agree with the decision to retract the paper with the exception of X.R., who was unable to be contacted.

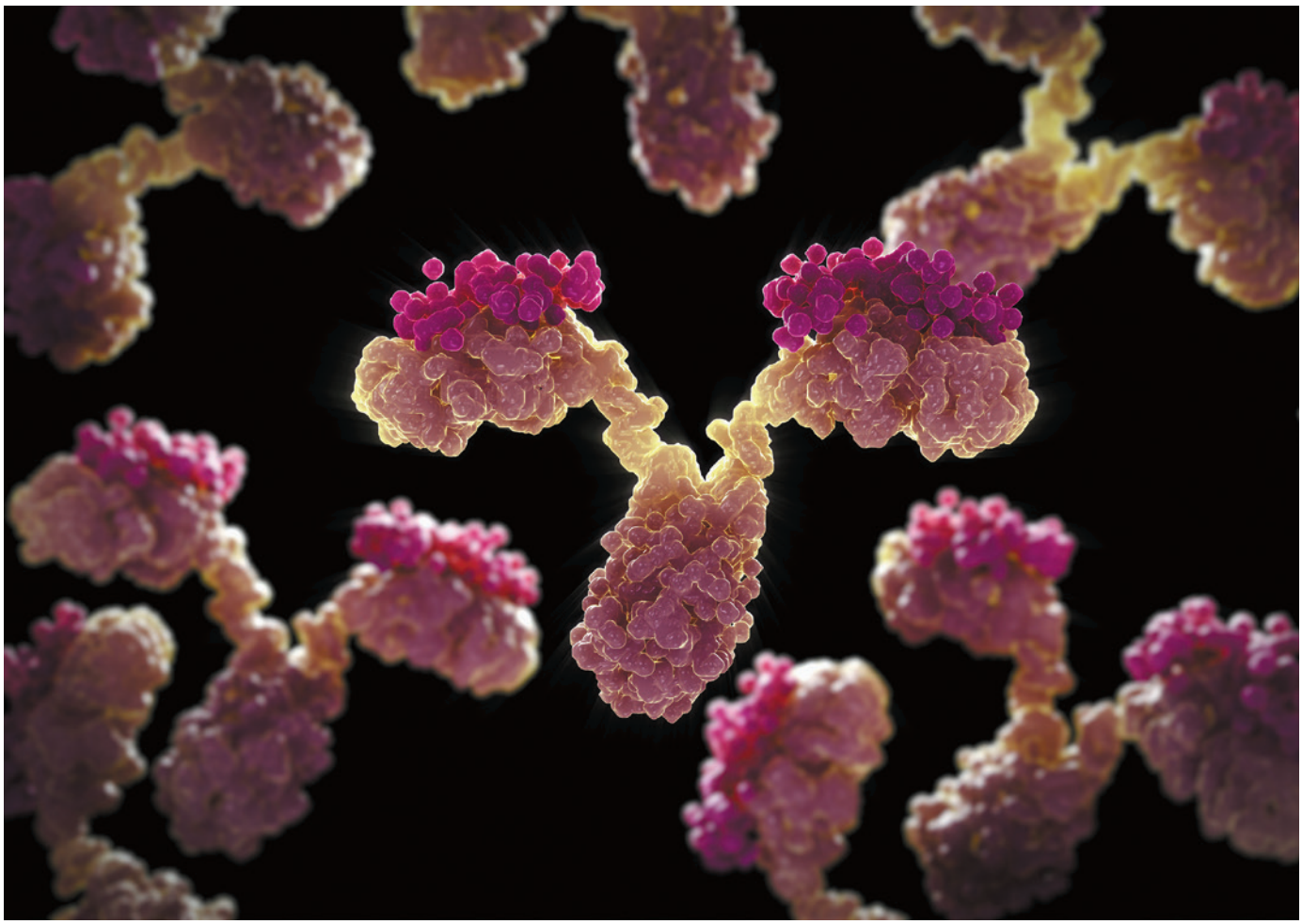
1. Rabouw, F. *et al.* Non-blinking single-photon emitters in silica. Preprint at: http://arXiv.org/abs/1509.07262 (2015).

TECHNOLOGY FEATURE

ANTIBODY ANARCHY: A CALL TO ORDER

Antibodies used in research often give murky results. Broader awareness and advanced technologies promise clarity.

SCIENCE PICTURE CO./SPL



Antibodies, with their distinctive Y-shape, are among the most widely used — and most vexing — reagents in biology.

BY MONYA BAKER

A mouse first alerted Clifford Saper to the fact that antibodies were misleading the scientific community. As editor-in-chief of the *Journal of Comparative Neurology* between 1994 and 2011, he handled scores of papers in which scientists relied on antibodies to flag the locations of neurotransmitters and their receptors. Around the turn of the century, related investigations began to roll in from researchers using knockout mice, animals genetically engineered to not express a target gene. The results were unsettling. Antibody staining in knockout animals should have shown radically different patterns from those in unmodified animals. But all too often the images were identical. "As we saw more and more retractions due to this, I began to realize that we had no systematic way to evaluate papers that used antibodies," recalls Saper, now chair of neurology at Beth Israel Deaconess Medical Center in Boston, Massachusetts.

Thus began a one-journal revolution. Saper and his editorial colleagues set up a policy of requiring extensive validation data on each antibody[1]. The policy was good for rigour, but not submissions, he recalls. "Many authors were caught in the middle, and found it easier to publish their papers elsewhere." But Saper persisted. His efforts eventually culminated in the JCN Antibody Database, an inventory of a few thousand antibodies that can be ▶

trusted for neuroanatomy.

Today, biomedical researchers still collect tales of antibody woe faster than country-music labels spin out sad songs. The most common grumble is the cheating reagent: the antibody purchased to detect protein X surreptitiously binds protein Y (and perhaps ignores X altogether). Another complaint is 'lost treasure': a run of promising experiments that stalls when a new batch of antibodies fails to reproduce previous findings (see 'A market in a bind').

But technological advances and shifts in the scientific community now promise to cut through this antibody quagmire.

Antibodies are ubiquitous tools in the life sciences. Perhaps their most popular use is in western blotting to reveal the presence of a particular protein in cells or tissue samples, but they are also used to visualize proteins under the microscope by immunohistochemistry and immunofluorescence, as well as in many other applications that stem from an antibody's presumed ability to bind specific biomolecules. A 2015 report from online purchasing portal Biocompare puts the market for research antibodies at US$2.5 billion a year and growing. The choice is dazzling: there are hundreds of vendors supplying products.

It is alarming, then, to discover that antibodies can be unreliable reagents. Insufficient specificity, sensitivity and lot-to-lot consistency have resulted in false findings and wasted efforts. Antibody unreliability has taken its toll across studies in cancer, metabolism, ageing, immunology and cell signalling, and in any field concerned with researching complex biomolecules. The waste, in terms of time and resources, is colossal. Losses from purchasing poorly characterized antibodies have been estimated at $800 million per year, not counting the impact of false conclusions, uninterpretable (or misinterpreted) experiments, wasted patient samples and fruitless research time[2].

Mathias Uhlén, a protein researcher at the Royal Institute of Technology in Stockholm, says that frustration with research antibodies has been building for years[3] and that the time is finally ripe for improvements. "There is a big interest in the community to clean this up."

SPURRED TO ACT

Discontent has spurred action along various fronts. In September, Uhlén chaired the inaugural meeting for a working group on antibody validation hosted by the Human Proteome Organization, an international consortium based in Vancouver, Canada, that supports large-scale projects for understanding proteins. That same month, the Federation of American Societies for Experimental Biology hosted roundtables to explore problems with antibodies. It expects to issue recommendations early next year. The US National Institutes of Health (NIH) is also on the case. Starting in January next year, grant applications must include a new section describing efforts to authenticate antibodies and other key resources required for experiments. Far-reaching solutions are likely to be hammered out at a meeting hosted by the Global Biological Standards Institute next September. The gathering will be held in Asilomar, California, where scientists gathered 40 years ago to set cautionary approaches for using recombinant genetic technology to manipulate DNA.

"We're hoping that the community will come up with consensus guidelines," says Jon Lorsch, director of the US National Institute of General Medical Sciences in Bethesda, Maryland. That way, both grant applicants and reviewers will have resources to turn to when describing how they will authenticate their materials.

Such resources could take the form of a menu of broad-strokes criteria. "We are not talking about good and bad antibodies but antibodies that work in specific assays and specific context," says Uhlén. Evaluation categories might include knockdown and knockout approaches to reveal whether an antibody still binds even in the absence of the target protein. Another approach would be to tag a target protein with a fluorescent marker to reveal whether the antibody also binds untagged proteins. A third category could compare a new antibody with a well-characterized one. Finally, researchers could run the antibody and whatever it binds through a mass spectrometer to analyse bound molecules for the expected protein fragments.

Several vendors have announced their own characterization efforts, and new technologies are helping. Alan Hirzel, chief executive officer of Abcam, a life-sciences reagents provider in Cambridge, UK, says that to verify that its commercial antibodies perform as expected, the

A market in a bind

An antibody that performs differently across experiments can cause calamity. But the performance of these reagents is linked to how they are manufactured.

Polyclonal antibodies are made by collecting the blood of an animal immunized with the target antigen. Any particular lot will therefore only be available as long as the animal lives. To produce monoclonal antibodies, a host animal is immunized with the target protein or relevant portion of it, then the B lymphocytes that recognize and respond to that antigen are fused to a myeloma cell line that can be cultured indefinitely to produce the desired antibody.

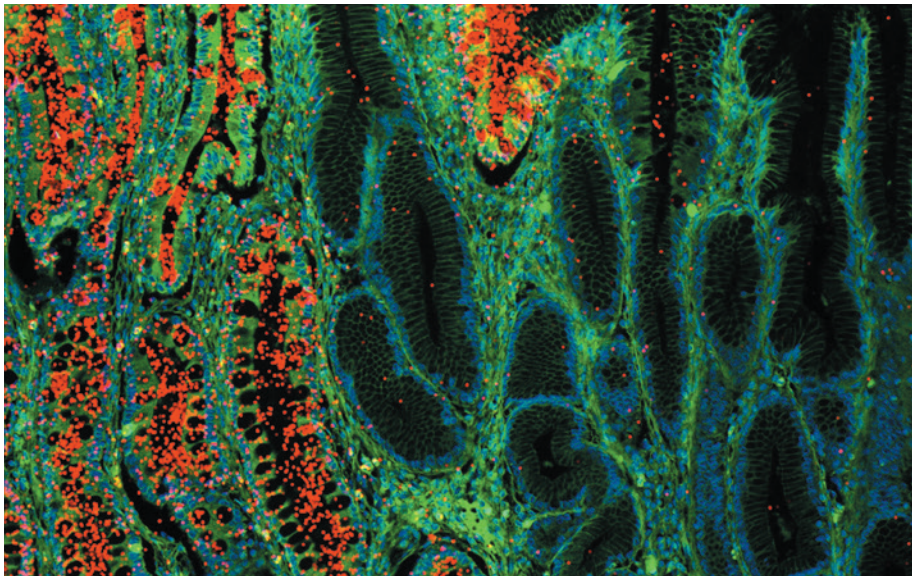
Recombinant antibodies are unlike traditional monoclonals because they can be manufactured without animals. Instead, these antibodies are made by identifying an exact gene sequence for an antibody — either by sequencing an animal's immune cells to find those that produce antibodies with highest affinity for the target, or sequentially shuffling gene sequences and testing the resultant proteins. That gene can then be introduced into an appropriate cell line to produce antibodies. Because the identity of the antibody is precisely defined, the cell line can be regenerated if the original colony dies or mutates.

The pursuit of antibody quality has inspired two publicly funded initiatives aimed at generating collections of validated antibodies and other protein-binding reagents. These produced thousands of new binders, but the Protein Capture Reagents programme, which launched in 2010, is already winding down, as is the European Union-funded Affinomics consortium, which launched in 2007 (ref. 8). Advocates say that the chosen targets, such as transcription factors, were particularly problematic and that further investments in such reagents would yield larger pay-offs.

Meanwhile, polyclonals command a large swathe of the market. A project that profiled reagents used across 10,000 biomedical papers published since 2006 found references to 1,293 polyclonals, 755 monoclonals and only 1 recombinant. Some researchers think that polyclonal antibodies, which can target a protein in multiple ways, are not only easy to manufacture but also particularly good at recognizing proteins in diverse contexts.

Eric McIntush is chief scientific officer of Bethyl Laboratories in Montgomery, Texas, which has been selling polyclonal antibodies for over 40 years and plans to start selling recombinants in 2016. The research world needs both, he says. Companies simply cannot afford to sink funds into products that they may never sell. The widespread availability of polyclonals, which are currently the least expensive antibody to develop, may encourage experiments on under-investigated proteins. As targets become more defined and are needed for translational applications, he says, there will be a market for recombinant products.

But researchers such as Andreas Plückthun, a protein engineer at the University of Zurich in Switzerland, think that polyclonals and monoclonals should be eliminated entirely in favour of defined binders. He agrees that many proteins are not addressed by existing reagents but does not see the point in making undefined products such as polyclonals. "Why not use something where the genes can be identified or kept?" he asks. M.B.



Pairs of antibodies can be designed to signal (red) only when both detect the same target protein[9].

company is using a genome editing method called CRISPR–Cas9, which makes precise changes in DNA. The company is testing antibodies on human cell lines in which target genes have been disrupted by CRISPR–Cas9 and then posting results for each reagent tested.

"We now really have the technologies we need that allow us to carry out those characterizations, whereas 5 or 10 years ago, we simply didn't," says Klaus Lindpaintner, chief scientific officer at Thermo Fisher Scientific, a life-sciences tools provider in Waltham, Massachusetts. Those companies with characterization data are starting to view this as a competitive advantage. In June this year, life-sciences company Bio-Rad in Hercules, California, launched a line of antibodies that have been tested for off-target activity in western blots against 12 different cell lines. Since mid-2014, Proteintech, an antibody manufacturer in Chicago, Illinois, has been using small interfering RNA to knock down gene expression in each new antibody product — assessing whether the signal subsides with the expression of the target gene. Such efforts are nascent, however, with only a tiny fraction of companies' catalogues being subjected to validation.

"Providers cannot guarantee that a given antibody will work for every tissue type."

And not all companies disclose the specific conditions of testing, or whether an antibody has performed poorly under those conditions, says Gordon Whiteley, lab director at the NIH's Antibody Characterization Program, which aims to create reliable antibodies for use in cancer biology. The example his programme sets in terms of supplying testing protocols and resulting data could be just as important as the reagents themselves, he says.

There will be no single best way to test antibodies, says Roberto Polakiewicz, chief scientific officer of Cell Signaling Technology, an antibody manufacturer in Danvers, Massachusetts. "Developing an antibody is a scientific endeavour. You need people who know what experiments to do to validate an antibody." If customers cannot see the data and make their own judgements, they need to look for a new antibody, he says.

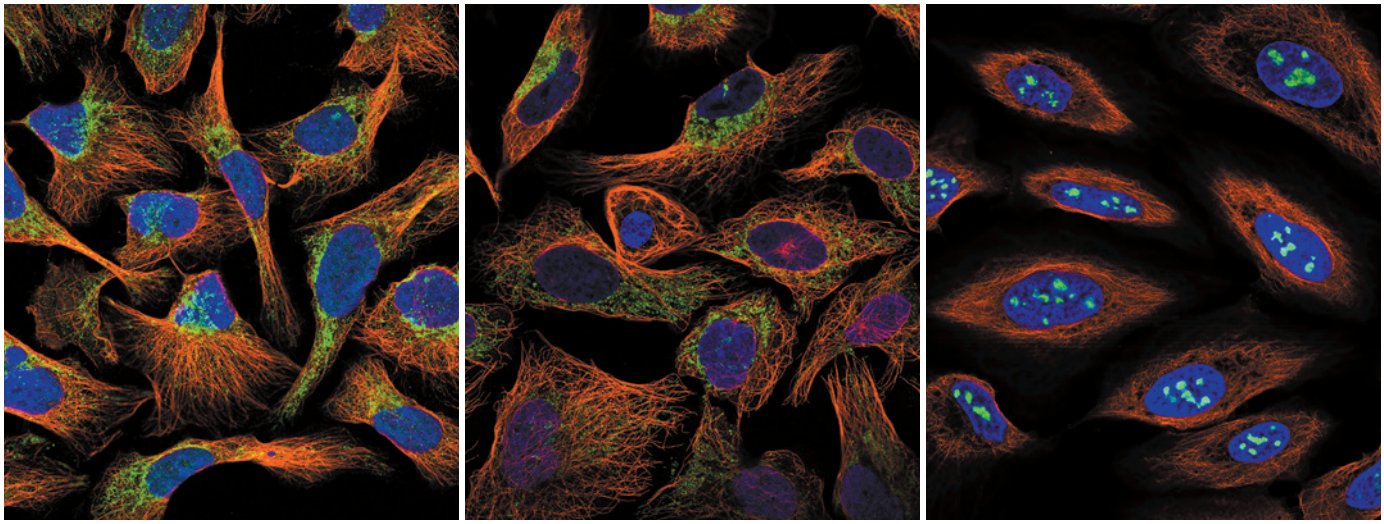
But researchers sometimes take only a cursory look at data, and many do not realize that antibodies' performance in a given tissue or application, such as western blotting, says little about whether it will work in other sorts of experiments.

And commercial providers cannot guarantee that a given antibody will work for every tissue type and experimental condition, warns Paul Sawchenko, a neuroscientist at the Salk Institute in San Diego, California. "Unless one is so fortunate as to have had someone else demonstrate specificity in the same tissue from the same species under the same experimental conditions, you should be obliged to do this yourself."

VITAL INFORMATION

It would be more efficient to learn from other researchers' work, but fewer than half of the publications that describe antibody experiments report which specific reagent was actually used[4]. Even when authors do include a catalogue number, companies may discontinue products and sell off lines, making them hard to track, says Anita Bandrowski, an information scientist at the University of California, San Diego. Bandrowski is group leader at the Resource Identification Initiative, an NIH-backed programme involving a diverse group of academic collaborators. The initiative has been instrumental in establishing unique identifiers for antibodies and persuading dozens of journals to ask authors to specifically name which antibodies they are using.

HUMAN PROTEIN ATLAS, WWW.PROTEINATLAS.ORG



Three antibodies (green) against the same mitochondrial protein. The unexpected pattern on the right shows the third antibody binds an unintended protein.

Information is beginning to accumulate. More than two dozen web portals have sprung up to help researchers select antibodies. Some collect user reviews on antibody performance and offer comparison tools. The Antibody Validation Channel, a project of the scientific publisher F1000, allows researchers to post their accounts and even request peer review. Biocompare has hired a content editor whose sole focus is to reach out to the research community and get them to write reviews.

Some antibody suppliers, such as St John's Laboratory in London, offer researchers free products in exchange for testing and sharing the results. Antibodies-online, a market place for antibodies, arranges for an independent third party to perform validation. At Antibodypedia's knockdown initiative, launched in September, life scientists can earn hundreds of dollars in free reagents if they submit data showing that gene-silencing reagents such as small interfering RNA or CRISPR–Cas9 eliminate an antibody signal for a given target.

But many scientists are wary of information from anonymous reviews. Data supplied by both users and companies can be sparse, and some projects share data only if they confirm that an antibody works as expected. "Sometimes it seems easier to hire a detective than to order a specific antibody," concludes an overview of antibody portals[5].

FUTURE ASSESSMENTS

Some researchers are developing mechanisms to compare antibodies directly. Aled Edwards at the University of Toronto, Canada, is director of the international Structural Genomics Consortium (SGC). He and his SGC colleagues used mass spectrometry to detect and compare the sets of proteins pulled down by immunoprecipitation with more than 1,000 antibodies[6]. The collaboration ran across 5 reference laboratories, took 4 years and cost US$3 million, not counting in-kind donations. Ultimately, it established a procedure to score antibody quality and share quantitative information about its performance, specifically for 'pull-down experiments', in which proteins are pulled out of solution using antibodies.

Fridtjof Lund-Johansen, a proteomics researcher at Oslo University Hospital in Norway, is developing an ambitious bead assay that tests thousands of antibodies at once[7]. The plan is to separate cellular proteins into many different fractions, then profile the proteins in each fraction using two different methods. One is mass spectrometry and the other is a bead-based array with thousands of antibodies. The mass spectrometry data serve as a reference for the results obtained with antibodies. Turning the idea into a refined assay will take considerable work, Lund-Johansen admits. "It is extremely ambitious. It is totally crazy, but it is the only way to go." Other scientists are intrigued at the approach but wonder if it will predict antibody performance in common techniques.

Blanket assessments of antibodies can be overinterpreted, says Ulf Landegren, a proteomics technology developer at Uppsala University in Sweden. "It is far more meaningful to discuss the ability of assays to detect the correct protein, rather than whether antibodies or other binders bind the right protein." A case in point is cross-reactivity, when an antibody binds proteins other than its specified target. Cross-reactivity depends not just on a particular antibody, but also on the complexity of a sample, the concentration of the antibody and the rarity of the target protein. He recommends that rather than relying on a single antibody, researchers should instead test antibodies in pairs that are designed to bind to different parts of a target protein. Parts of a sample labelled with both reagents are less likely to represent off-target binding.

One problem with this approach is that it is hard for scientists to know if they are purchasing different antibodies. Vendors often obtain products from different sources and are not required to disclose the original manufacturer. As a result, researchers who want to compare several antibodies may end up comparing identical products sold by several vendors. A handful of companies, including Genlogica and One World Laboratories, both in San Diego, California, only sell products labelled by the original manufacturer and offer 'trial size' antibody batches so that researchers can test products side by side in their labs.

The toughest challenge is not so much in antibody characterization but in persuading cell biologists to hold back on using antibodies until these are thoroughly evaluated, says Edwards, although he doubts that scientists will become savvier unless funders and publishers force the issue. "Right now we have an unregulated market, where you don't have to have any quality to sell your product." In other words, he says, guidelines, characterization data and conscientious vendors only matter if researchers invest effort into selecting reagents. ■

Monya Baker *writes and edits for* Nature *in San Francisco, California.*

1. Saper, C. B. & Sawchenko, P. E. *J. Comp. Neurol.* **465,** 161–163 (2003).
2. Bradbury, A. *et al. Nature* **518,** 27–29 (2015).
3. Bordeaux, J. *et al. BioTechniques* **48,** 197–209 (2010).
4. Vasilevsky, N. A. *et al. PeerJ* **1,** e148 (2013).
5. Pauly, D. & Hanack, K. *F1000Research* http://dx.doi.org/10.12688/f1000research.6894.1 (2015).
6. Marcon, E. *et al. Nature Methods* **12,** 725–731 (2015).
7. Wu, W. *et al. Mol. Cell. Proteomics* **8,** 245–257 (2009).
8. *Nature Methods* **12,** 373 (2015).
9. Conze, T. *et al. Glycobiology* **20,** 199–206 (2010).

CORRECTION

The Technology Feature 'Connectomes make the map' (*Nature* **526,** 147–149; 2015) misnamed the MultiSEM model and gave the wrong citation in reference 3. MultiSEM 505 should have been Zeiss MultiSEM, and ref. 3 should have referred to Zingg, B. *et al. Cell* **156,** 1096–1111 (2014).

CLAIMED

First contact.

BY S. J. ROSENSTEIN

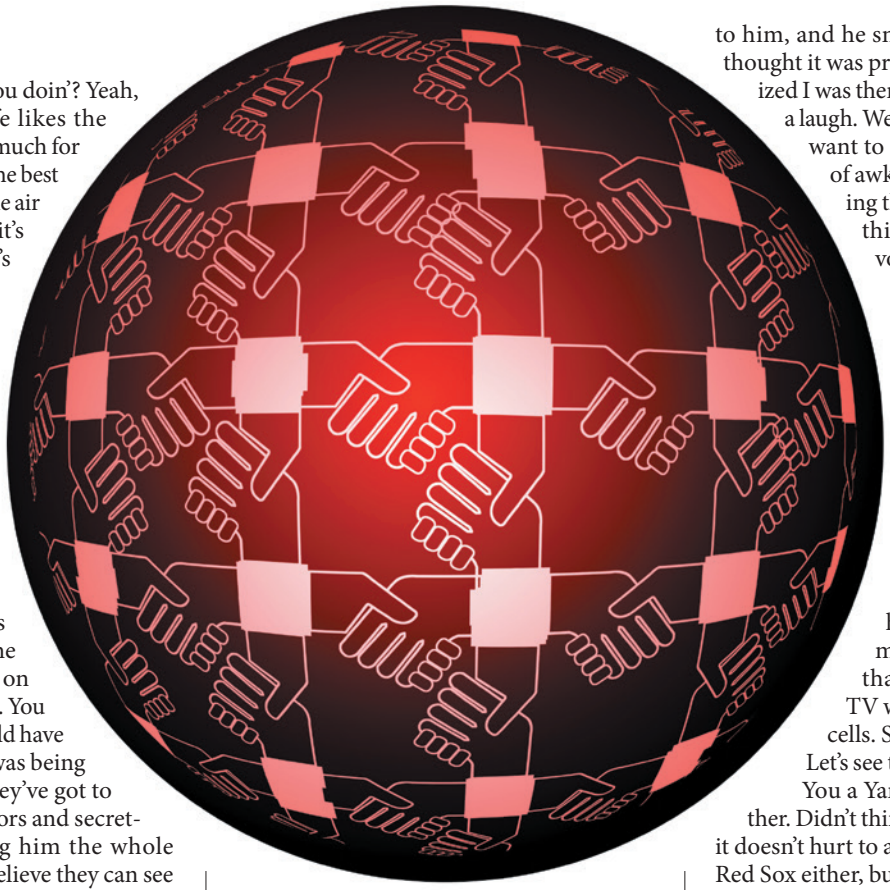


ILLUSTRATION BY JACEY

Hey, there. How you doin'? Yeah, me too. My wife likes the heat, but it's too much for me. It's why I come here, the best thing about this place is the air con. No one would say it's the coffee. No offence. It's good enough for me, I'm happy as long as I don't have to drink any of that fancy frou-frou stuff. I'll have mine black. My wife made me give up cream, she won't stop going on about my cholesterol.

Yeah, I seen the news. Come on now, you don't believe that garbage, do you? Them Claimers is as crazy as my mom, and she thinks the CIA is spying on her through her sprinkler. You don't think someone would have noticed if the president was being controlled by aliens? They've got to have about a dozen doctors and secret-service agents watching him the whole time. Don't tell me you believe they can see the future, too. That's what they say. Kooks and slackers with nothing better to do with their lives, sitting in their basements working themselves into a frenzy over nothing. Look here now, if I was an alien and I could see the future, I'd just buy me a lottery ticket and retire to Hawaii.

No, you're just plain wrong. I got more right to an opinion than you, 'cos I know what I'm talking about. You think just 'cos I don't wear a thousand-dollar suit I've never met the great and the good? I met the president only two weeks ago. Yeah. Yeah, I guess it was exciting. Well, the company I work for, they make the toilets for the shuttle. The one the president used to go up to the alien ship, yeah. Well, just on the way back, after he'd met the aliens and taken all those photos you seen in the papers, the toilet got blocked up. They didn't know how soon they'd need to go back, so I drove up right away. Blew through a few stop signs too, I kind of hoped a cop would stop me so I could tell them I was on the way to fix the president's toilet. Didn't happen, though.

So I got to the spaceport just as the shuttle was getting in, and the president got off. I'd sort of wormed my way forward so I could get on as fast as possible, 'cos they'd been pretty mad about the toilet and I wanted to show that even though we're just a tiny company, we still give a great service. The president looked real tired, and as I was trying to get to the door he looked up and saw me, and took a few steps towards me, and shook my hand. This hand right here touched the president. Huh. No, I dunno what that is. Little scratch or something. Well, it looks a bit weird, but it doesn't feel infected. It's been there a couple of weeks. Nothing to worry about.

Anyway, I think the president thought I was a foreign dignitary or something, he started talking about the future of our world and strategic realignments and trade agreements, and what a great deal the aliens were offering us and how wonderful the future was gonna be. I just sort of stood there, didn't really know what to say. One of the spooks got to him pretty quick and whispered something to him, and he snapped out of it OK. He thought it was pretty funny when he realized I was there to fix the toilet. We had a laugh. Well, no, I didn't. But I didn't want to tell him. Would be kind of awkward, wouldn't it, meeting the president and the first thing you say is you didn't vote for him?

So I went in and fixed the toilet. Wouldn't you know it, one of the spooks had managed to drop one so big and dense it blocked the system. No, I don't think it was the president. He just doesn't look like that sort of guy. Anyways, there you have it. I've met him, and so I think I have more right to an opinion than some woman on the TV with more hair than brain cells. Switch it over, won't you? Let's see the game.

You a Yankees fan? Nah, me neither. Didn't think so, not down here, but it doesn't hurt to ask. Not like I'd cheer the Red Sox either, but you gotta feel sorry for their fans. Mind you, they'll be happy this time. Well, I know it looks like that, but they ain't gonna lose. That guy coming up to the plate, he'll hit a home run. Don't ask me how I know, I just know. Same way you wake up in the morning sometimes and you know it's gonna be the sort of day when a bird craps on you. Only, like, stronger. There he goes. See, didn't I tell you? You should have put money on it.

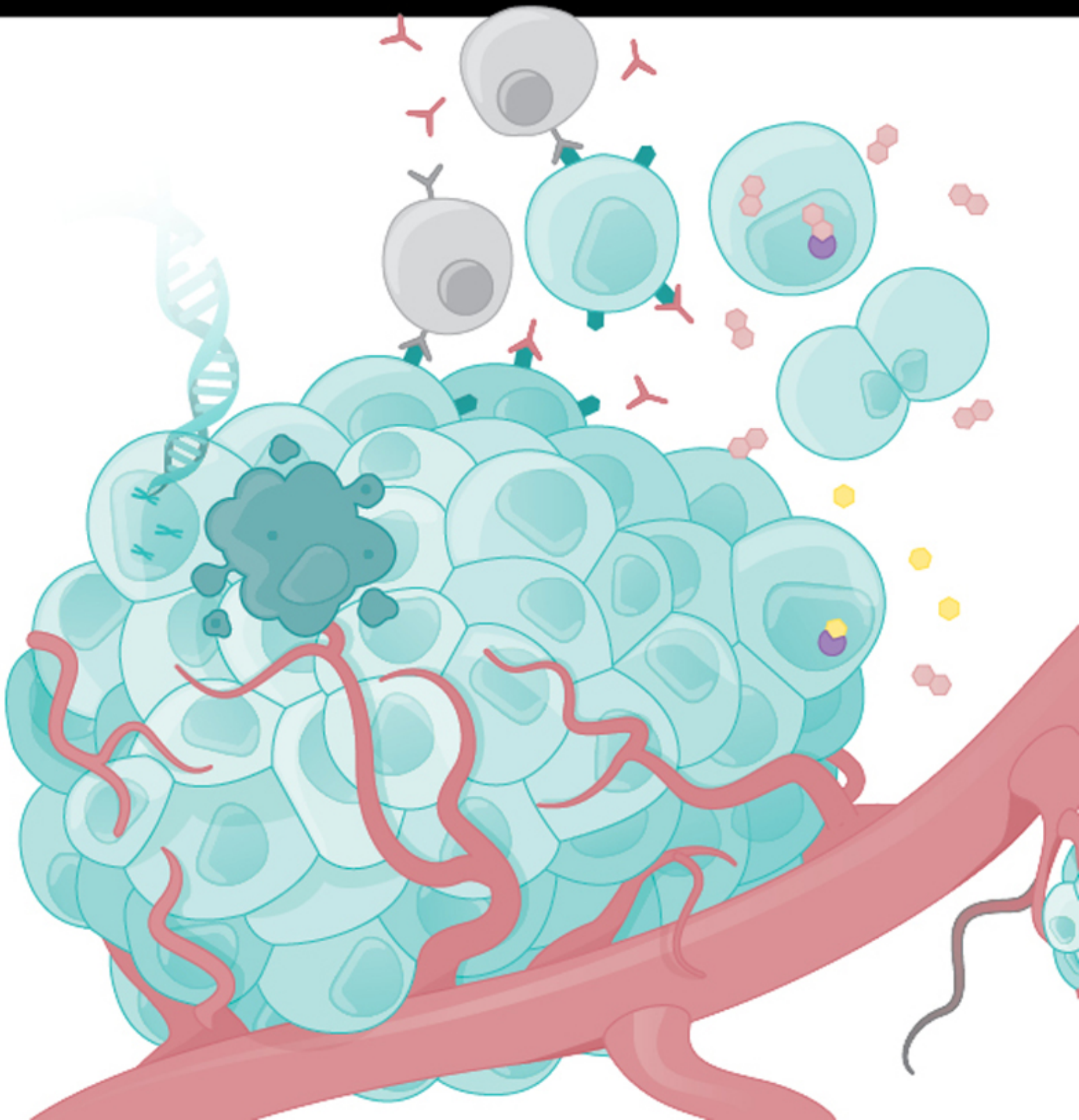
Well, I got to be going. Look here, no hard feelings, eh? Sorry I was a bit touchy, I just get so mad when I hear the nonsense people are spouting. Shake on it? Oh, I'm sorry. Must have scratched you with my ring or something. Don't worry, it'll heal right up. Well thanks, I'm glad to hear you say that. It sure makes me feel good, knowing someone like me can change the way someone else thinks. You have a good day too. ■

S. J. Rosenstein *is a research scientist with a secret identity as a writer, although both incarnations wear glasses and neither are particularly mild mannered. She complains about life at alackoftheologyandgeometry.wordpress.com.*

NATURE.COM
Follow Futures:
@NatureFutures
go.nature.com/mtoodm

nature OUTLINE

OVARIAN CANCER: BEYOND RESISTANCE



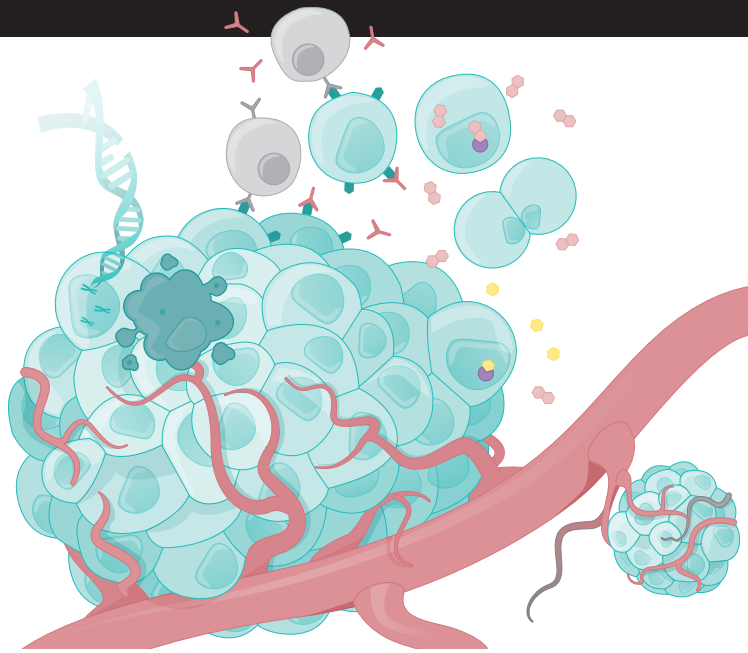
nature OUTLINE

Watch an animation at: go.nature.com/ghn2pe | Infographic: Pages S218–S219

OVARIAN CANCER: BEYOND RESISTANCE

Successfully treating the cancer requires overcoming the almost inevitable development of resistance to standard platinum-based therapy.

BY DAVID HOLMES



LUCY READING-IKKANDA

Ovarian cancer is the most common cause of gynaecological-cancer-associated death. Although the past 40 years have seen our knowledge of the disease advance, translating that improved understanding into a tangible clinical benefit has been a tortuous process. The last major breakthrough in treatment came 20 years ago, with the addition of a taxane (paclitaxel or docetaxel are commonly used to treat ovarian cancer) to one of the several variants of platinum-based chemotherapy that remain the mainstay of treatment. Since then, refinements to surgery and to the timing and delivery of chemotherapy have produced only slight improvements in outcomes. In the United States, for example, 5-year survival has inched up from about 40% in 1985 to a still parlous 45% today. By comparison, 5-year survival for breast cancer stands at 90%.

Two factors account for much of the stubbornly high mortality and morbidity associated with ovarian cancer — late diagnosis and treatment resistance. There are currently no approved methods to screen for ovarian cancer, although promising preliminary results released earlier this year from the UK Collaborative Trial of Ovarian Cancer Screening may begin to change that. Around 60% of women are diagnosed with late-stage disease that has already spread within the abdomen. As many as 80% of these women will respond well to initial treatment with platinum-based therapy, but almost all will experience multiple recurrences of disease, with ever shorter disease-free intervals. Ultimately, almost all of these women will die from the disease — and most will die from a disease that is resistant to platinum chemotherapy.

As with most solid malignancies, resistance to platinum-based treatment can be intrinsic or acquired, and is brought about through a bewildering array of mechanisms. From pumps that eject the drug from the cell to promoting the expression of genes that enable alternative growth pathways, cancer cells leave no stone unturned in their bid to survive and proliferate. Further complicating matters is the difficulty of knowing which mechanism or mechanisms are active in any particular person.

The good news is that there are a huge number of experimental therapies in development, and it is hoped that these can be added to platinum-based chemotherapy to help deliver a knockout blow, or at least prolong the intervals between treatment and improve patients' quality of life. Vaccines to activate the immune system against tumours, agents to interfere in DNA-repair pathways, and therapies that choke off the supply of blood to the tumour are all now in clinical trials for ovarian cancer. Because of the complexity of the disease and the mechanisms that underlie treatment resistance, it is unlikely that any one therapy will be a silver bullet. Nevertheless, there is a growing sense of optimism that researchers will be able to translate hard-won knowledge into improved outcomes for patients.

Nature is pleased to acknowledge the financial support of Pharma Mar, S.A. in producing this Outline. As always, *Nature* retains sole responsibility for all editorial content. ■

David Holmes *is a science writer based in the United Kingdom.*

Produced with support from: Pharma Mar

THE PROBLEM WITH PLATINUM

For an animated version of this graphic visit: go.nature.com/ghn2pe

Ovarian cancer is difficult to treat, largely because tumours are often found late and develop resistance to initial treatment: platinum-based therapy. New approaches promise to break through the platinum barrier. By **David Holmes**; *illustration by* **Lucy Reading-Ikkanda.**

BIG PROBLEM, LITTLE PROGRESS

BAD TIMING

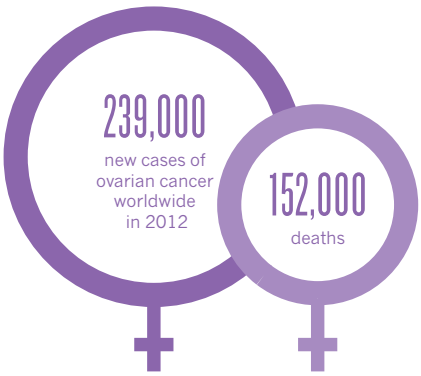
The earlier that ovarian cancer is identified, the better the odds are that treatment will be successful. Women are not screened because current methods are not reliable enough to predict whether or not women have the disease. Early symptoms of ovarian cancer are often confused with irritable bowel or premenstrual syndrome, so most people are diagnosed with late-stage disease.

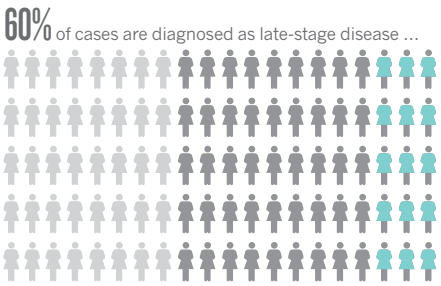
THE TOLL OF RESISTANCE

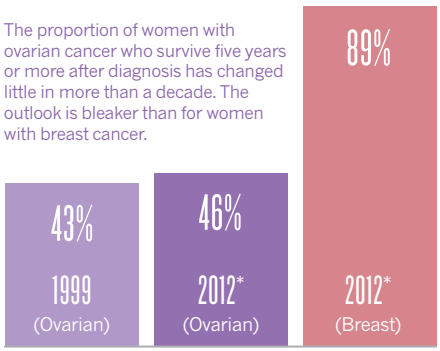
Of US women diagnosed with ovarian cancer, 60% have late-stage disease. Most of these initially respond well to treatment with a combination of paclitaxel (a drug that interferes with cell division) and carboplatin (a platinum-based drug that damages cancer-cell DNA). However, more than half will relapse within 18 months of diagnosis.

SLOW PROGRESS

The late stage at which most ovarian cancers are diagnosed, the fact that such a high proportion become resistant to platinum-based chemotherapy, and the small number of approved alternatives to platinum therapy, mean that ovarian cancer has a relatively low five-year survival rate. In the United States, for example, it is just 45.6%.

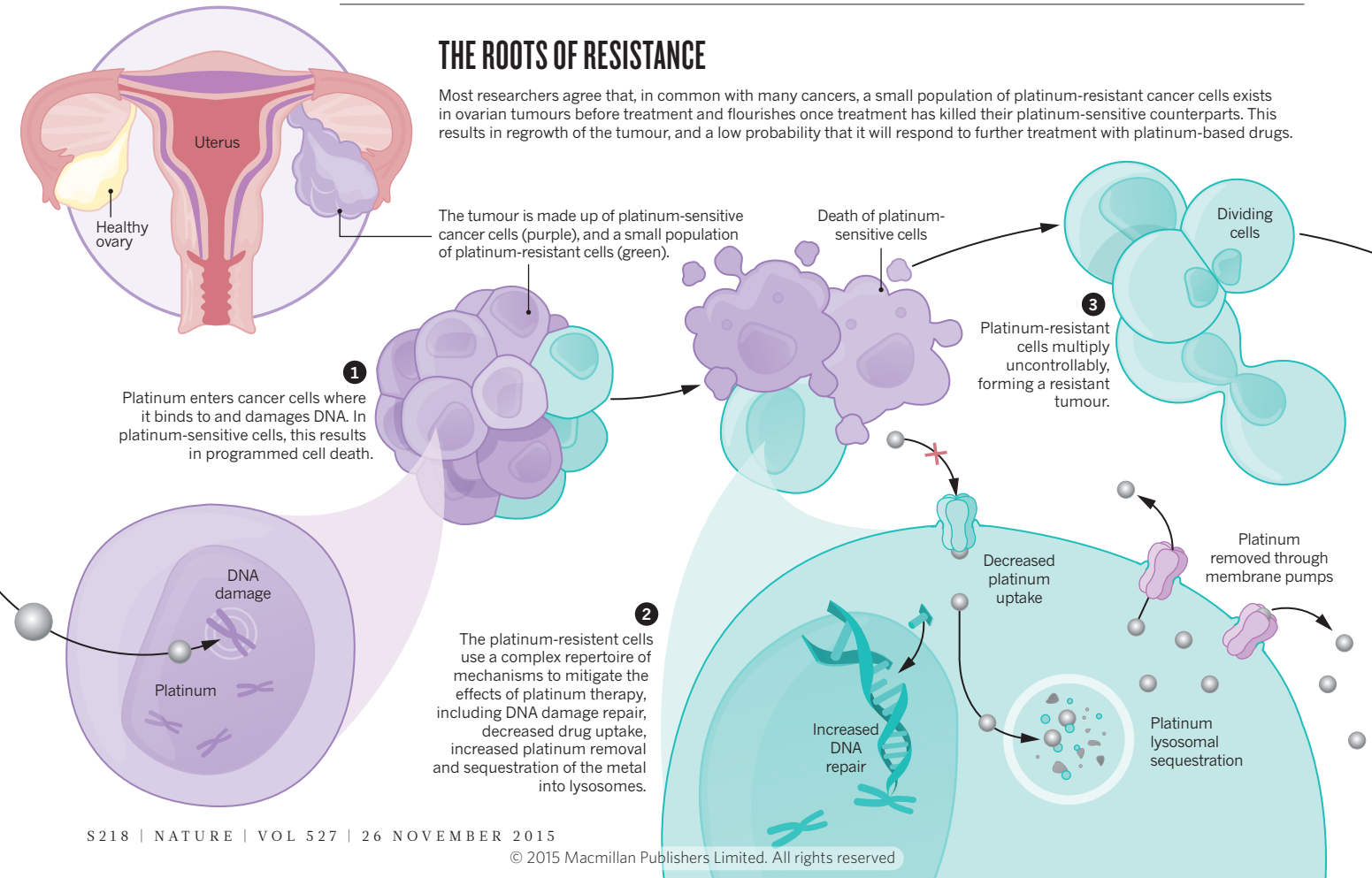






THE ROOTS OF RESISTANCE

Most researchers agree that, in common with many cancers, a small population of platinum-resistant cancer cells exists in ovarian tumours before treatment and flourishes once treatment has killed their platinum-sensitive counterparts. This results in regrowth of the tumour, and a low probability that it will respond to further treatment with platinum-based drugs.



THE NEW WAVE OF THERAPY

From priming the immune system to fight ovarian tumours to cutting off the cancer's blood supply, researchers are testing a variety of ways to overcome resistance to platinum-based chemotherapy.

SCRAMBLING THE CODE

Ramping up DNA-repair pathways is one of the ways that cancer cells resist the DNA-damaging effects of platinum. If those DNA-repair pathways could be dampened down it might be possible to resensitize cancer cells to platinum. There are several drugs in development that aim to do just that. PARP inhibitors disrupt the mechanism by which damaged parts of DNA are removed, and the drug trabectedin binds directly to and damages the DNA. Both have shown early promise. The drug topotecan blocks the action of the enzyme TOP1, which helps to repair DNA damage, and is already licensed for the treatment of recurrent ovarian cancer. However, its effect on overall survival is limited.

DNA damage and disruption of DNA-repair mechanisms leads to cell death.

IMMUNE BOOSTERS

Priming the immune system to recognize and attack cancer cells might be an effective way of stunting the growth of tumours in people with recurrent ovarian cancer. A UK trial called TRIOC is testing whether the TroVax vaccine, which has this priming effect, can boost an individual's anticancer immune response enough to slow the growth of recurrent ovarian tumours and delay the need for a second line of chemotherapy. In the trial, the vaccine is given to people who have high levels of a marker called CA125 in their blood, which indicates that a cancer may have returned.

The vaccine-primed immune system releases antibodies and T cells to bind to antigens on the cell surface.

HORMONE THERAPY

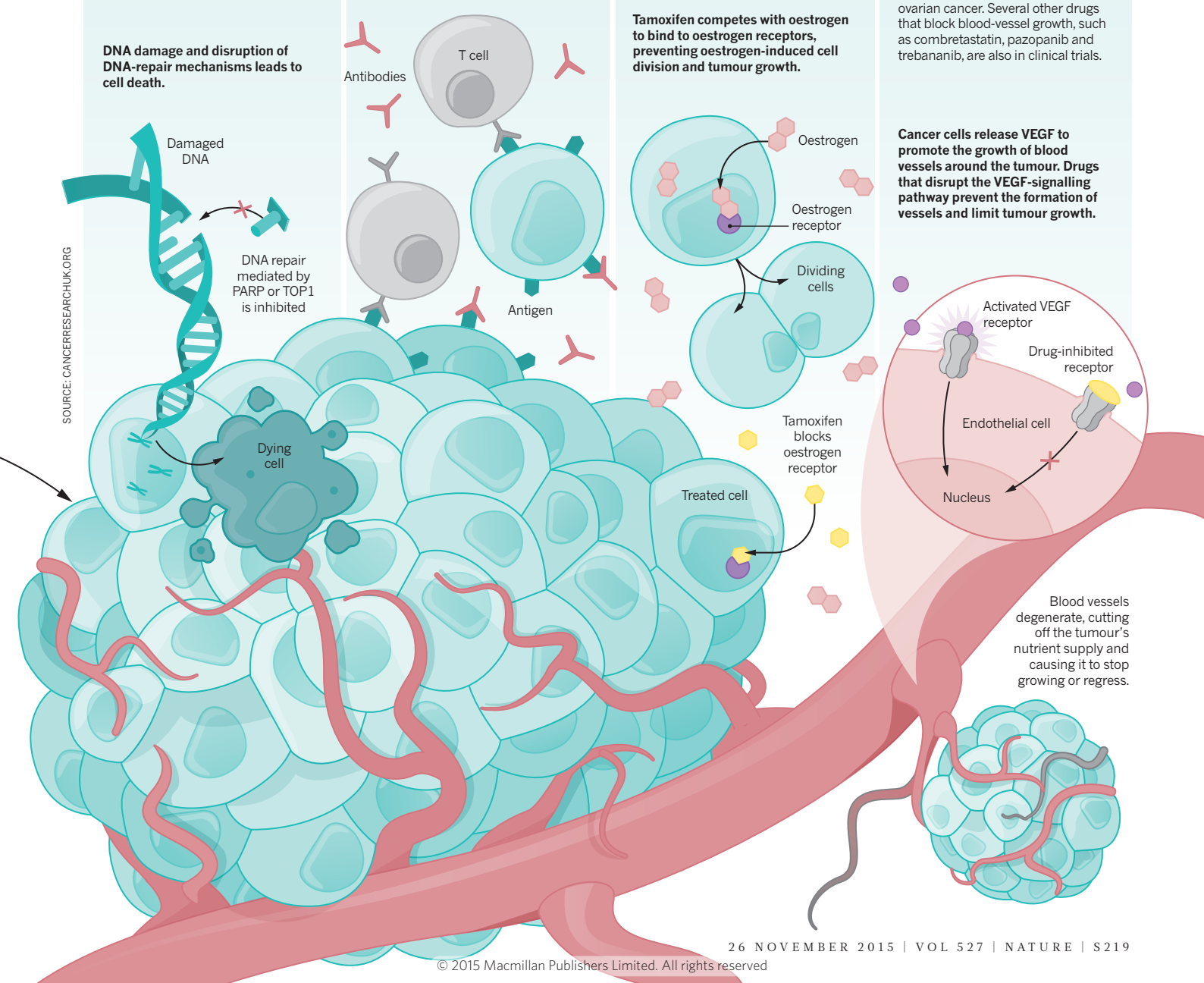
Similar to many breast cancers, some ovarian cancer cells have oestrogen receptors on their surface and may require the hormone to grow and spread. This has led researchers to test the hormone treatment tamoxifen, which is often used to treat oestrogen-receptor-positive breast cancers, in women with advanced ovarian cancer. Tamoxifen blocks oestrogen from reaching the cells and has been shown to work for a small proportion of women with recurrent cancer that does not respond to chemotherapy. Several other hormone treatments, such as letrozole and anastrozole, are also in clinical trials.

Tamoxifen competes with oestrogen to bind to oestrogen receptors, preventing oestrogen-induced cell division and tumour growth.

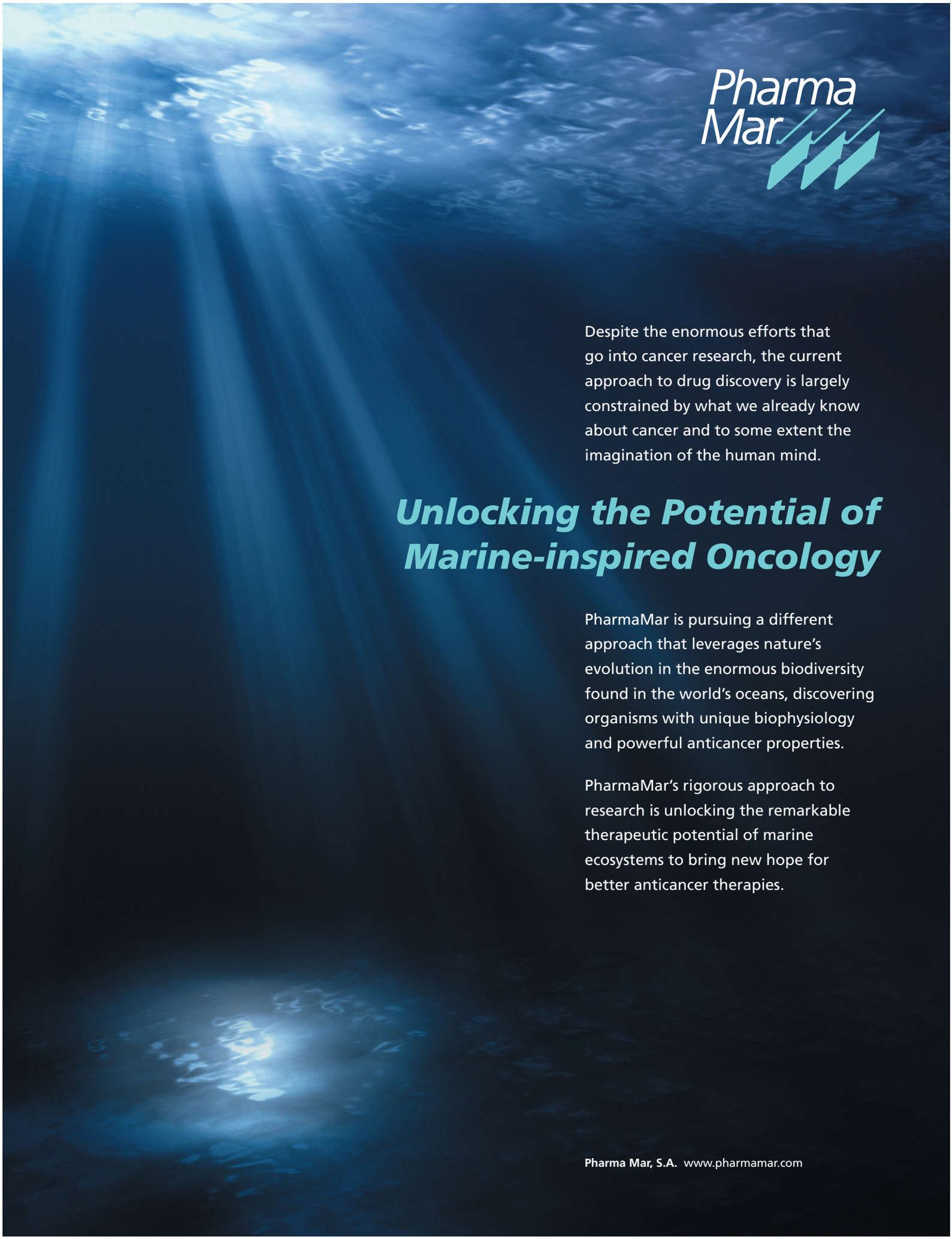
STARVING THE TUMOUR

Several treatments are in clinical trials to assess whether blocking the blood supply to tumours can slow down their recurrence. The antibody bevacizumab prevents the formation of new blood cells by inhibiting activity of the signalling protein VEGF, which is involved in the growth of blood vessels. The drug has already been approved by the US Food and Drug Administration and the European Medicines Agency for use in combination with chemotherapy for platinum-resistant relapsed ovarian cancer. Another drug — cediranib — disrupts the formation of blood vessels around the tumour by inhibiting a type of signalling protein called tyrosine kinase. In a trial called ICON6, the drug increased survival by three months compared with standard treatment for recurrent ovarian cancer. Several other drugs that block blood-vessel growth, such as combretastatin, pazopanib and trebananib, are also in clinical trials.

Cancer cells release VEGF to promote the growth of blood vessels around the tumour. Drugs that disrupt the VEGF-signalling pathway prevent the formation of vessels and limit tumour growth.



SOURCE: CANCERRESEARCHUK.ORG



Despite the enormous efforts that go into cancer research, the current approach to drug discovery is largely constrained by what we already know about cancer and to some extent the imagination of the human mind.

Unlocking the Potential of Marine-inspired Oncology

PharmaMar is pursuing a different approach that leverages nature's evolution in the enormous biodiversity found in the world's oceans, discovering organisms with unique biophysiology and powerful anticancer properties.

PharmaMar's rigorous approach to research is unlocking the remarkable therapeutic potential of marine ecosystems to bring new hope for better anticancer therapies.

Pharma Mar, S.A. www.pharmamar.com